(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

CARL MURCHISON

GORDON ALLPORT
Harvard University

ARTHUR G. BILLS
University of Cincinnati

E. A. BOTT
University of Toronto

LEONARD CARMICHAEL
Smithsonian Institution

HERBERT S. CONRAD
U. S. Office of Education

ELMER K. CULLER
University of Rochester

JOHN F. DASHIELL
University of North Carolina

ROY M. DORCUS
University of California at Los Angeles

HORACE B. ENGLISH
Ohio State University

FRANKLIN FEARING
University of California at Los Angeles

FRANK A. GELDARD
University of Virginia

CLARENCE H. GRAHAM
Columbia University

ERNEST R. HILGARD
Stanford University

HEINRICH KLÜVER
University of Chicago

CARNEY LANDIS
Columbia University

NORMAN L. MUNN
Bowdoin University

JOHN PAUL NAFE
Washington University

DONALD G. PATERSON
University of Minnesota

LEE EDWARD TRAVIS
University of Southern California

ROBERT C. TRYON
University of California

MICHAEL J. ZIGLER
Wellesley College

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

VOLUME 52

1961

Copyright, 1961, by The Journal Press
Published Quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

AUTHOR INDEX

Abramson, Harold A. 223, 445
Anderson, M. L. 491
Archibald, Herbert C. 281
Barmack, Joseph E. 3
Bartley, S. Howard 137, 241, 457, 479
Becknell, James C. 47
Bell, Dorothy 281
Berman, Leo H. 157
Binder, Arnold 379
Bourassa, Charles M. 479
Bravo, Lucy 333
Cappone, Margaret Kathleen 327
Casella, Carmine 241
Clausen, Johs. 25
Collins, Beverly 369
Crowne, Douglas P. 161
Cunningham, Scott M. 77
D'Amato, M. R. 411
Davis, Roger T. 403
Dean, G. 445
Dillon, Donald J. 105
Eichbauer, John 423
England, Samuel J. M. 149
Epstein, William 41
Fillenbaum, Samuel 133
Fisichelli, Vincent R. 423
Forney, Robert B. 185
Freedman, Lawrence Zelic 157
Gengerelli, J. A. 253, 317
Gettner, H. H. 445
Gilmer, B. Von Haller 211
Glasscock, Wilford R. 317
Gogel, Walter C. 287, 303
Gray, Florence E. 363
Hall, E. R. 99
Halpern, Seymour 265
Hanson, John A. 363
Harleston, Bernard W. 77
Hewitt, M. P. 445
Hoffman, L. Richard 429
Holter, Norman J. 317
Hughes, Francis W. 185
Jagoda, H. 411
Johnson, D. M. 99
Jones, Frank Pierce 363
Jones, Joan E. 335
Jorgensen, Ronald L. 403
Kanfer, Frederick H. 115
Karelitz, Samuel 423
Karrer, Rathe 25
Liverant, Shephard 161
Mackavey, William R. 241
MacKinnon, William J. 347
Mayzner, M. S. 333
McBrearty, John F. 115
McCollough, Celeste 63
McConnell, R. A. 87, 491
McReynolds, Paul 369
Miller, Christine 281
Mitchell, John D. 269
Mitchell, Mildred B. 125
Nelson, Thomas M. 137, 457, 479
Noble, Clyde E. 201
Pasamanick, Benjamin 149
Payne, Donald E. 3
Ranney, Jane E. 137
Rosenfeld, Laura S. 423
Rotter, Julian B. 161
Rudin, Stanley A. 179
Salop, Phyllis 379
Schrier, Allan M. 193
Schwartz, E. K. 249
Scodel, Alvin 55
Thompson, Clare W. 281
Tuddenham, Read D. 231
Wilson, Warner R. 47
Zagona, Salvatore V. 347

TABLE OF CONTENTS

Injury-producing private motor vehicle accidents among airmen: psychological models of accident-generating processes 3
By Joseph E. Barmack and Donald E. Payne

Perceived rate of electrical phosphenes as determined by matching square wave and sine wave flicker rates 25
By Johs. Clausen and Rathe Karrer

The relation between the association value, pronunciability, and affectivity of nonsense syllables 47
By Warner R. Wilson and James C. Becknell

Phenomenal orientation and perceived achromatic color 51
By William Epstein

Value orientations and preference for a minimax strategy 55
By Alvin Scodel

Forming and acting on impressions 63
By Celeste Mc Collough

Task difficulty and anxiety level as factors affecting performance in a verbal learning situation 77
By Bernard W. Harleston and Scott M. Cunningham

The discontinuity in Schmeidler's ESP-Rorschach data 87
By R. A. McConnell

Organization of relevant and irrelevant words in the solution of verbal problems 99
By D. M. Johnson and E. R. Hall

The variation of choice reaction time among groups of hospitalized psychiatric patients 105
By Donald J. Dillon

Verbal conditioning: Discrimination and awareness 115
By Frederick H. Kanfer and John F. McBrearty

Trends toward multiple authorship in scientific publications 125
By Mildred B. Mitchell

How fat is fat? Some consequences of similarity between judge and judged object 133
By Samuel Fillenbaum

The sensory parallel of the reorganization period in the cortical response in intermittent retinal stimulation 137
By S. Howard Bartley, Thomas M. Nelson, and Jane E. Ranney

The geochronograph method for representing nonstationary electrograms, such as heartbeat and respiration 149
By Samuel J. M. England and Benjamin Pasamanick

Clinical perception of sexual deviates 157
By Leo H. Berman and Lawrence Zelic Freedman

The growth and extinction of expectancies in chance controlled and skilled tasks 161
By Julian B. Rotter, Shephard Liverant, and Douglas P. Crowne

The relationship between rational and irrational authoritarianism . . . 179
By Stanley A. Rudin

Delayed auditory feedback and ethanol: Effect on verbal and arithmetical performance 185
By Robert B. Forney and Francis W. Hughes

A modified version of the Wisconsin General Test Apparatus 193
By Allan M. Schrier

Meaningfulness (*m*) and transfer phenomena in serial verbal learning . . 201
By Clyde E. Noble

Toward cutaneous electro-pulse communication 211
By B. Von Haller Gilmer

Intractable asthma: Conflict of period of toilet training 223
By Harold A. Abramson

The influence of a distorted group norm upon judgments of adults and children 231
By Read D. Tuddenham

Measurements of simultaneous brightness contrast across the retina . . . 241
By William R. Mackavey, S. Howard Bartley, and Carmine Casella

The analysis of mutual concurrences 253
By J. A. Gengerelli

The role of hypnosis in self understanding 265
By Seymour Halpern

A psychoanalytic approach to Kabuki: a study in personality and culture . . 269
By John D. Mitchell and E. K. Schwartz

Psychosomatic V 281
By Herbert C. Archibald, Dorothy Bell, Christine Miller, and Clare W. Thompson

Convergence as a cue to absolute distance 287
By Walter C. Gogel

Convergence as a cue to the perceived distance of objects in a binocular configuration 303
By Walter C. Gogel

Magnetic fields accompanying transmission of nerve impulses in the frog's sciatic 317
By J. A. Gengerelli, Norman J. Holter, and Wilford R. Glasscock

The effect of analgesic dosages of acetylsalicylic acid upon psychological reaction time 327
By Margaret Kathleen Cappone

Assimilation and contrast effects of anchoring stimuli on judgments: a partial replication of the Sherif, Taub, and Hovland study 333
By Lucy Bravo and M. S. Mayzner

Deviations from matching behavior in probability learning 335
By Joan E. Jones

Open-mindedness and hypotheses of interdependence between libertarian and equalitarian processes 347
By Salvatore V. Zagona and William J. MacKinnon

Head balance and sitting posture II: the role of the sternomastoid muscle . 363
By Frank Pierce Jones, John A. Hanson, and Florence E. Gray

Concept-forming behavior in schizophrenic and non-schizophrenic subjects . 369
By Paul McReynolds and Beverly Collins

Reinforcement and personality factors in verbal conditioning 379
By Arnold Binder and Phyllis Salop

Response strength as a function of rate of reward 403
By Ronald L. Jorgensen and Roger T. Davis

Studies on transfer of brightness discrimination from bar-pressing to a Y maze 411
By M. R. D'Amato and H. Jagoda

Volume-unit graphs: their production and applicability in studies of infants' cries 423
By Vincent R. Fisichelli, Samuel Karelitz, John Eichbauer and Laura S. Rosenfeld

Conditions for creative problem solving 429
By L. Richard Hoffman

Lysergic acid diethylamide (LSD-25) XXXIII. Effect on surfacing behavior of the carp and possible application to game conservation and survival kits 445
By H. A. Abramson, H. H. Gettner, M. P. Hewitt, and G. Dean

The roll of PCF in temporal manipulations of color 457
By Thomas M. Nelson and S. Howard Bartley

The effect of areal characteristics of targets upon shape-slant invariance . . 479
By Thomas M. Nelson, S. Howard Bartley and Charles M. Bourassa

Fantasy testing for ESP in a fourth and fifth grade class 491
By M. L. Anderson and R. A. McConnell

$20.00 per annum
$15.00 per volume
$7.50 per half volume

QUARTERLY
Two volumes per year
Immediate publication

July, 1961
Volume 52, First Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

JULY, 1961

(Manuscripts are printed in the order of final acceptance)

Injury-producing private motor vehicle accidents among airmen: psychological models of accident-generating processes 3
By Joseph E. Barmack and Donald E. Payne

Perceived rate of electrical phosphenes as determined by matching square wave and sine wave flicker rates 25
By Johs. Clausen and Rathe Karrer

The relation between the association value, pronunciability, and affectivity of nonsense syllables 47
By Warner R. Wilson and James C. Becknell

Phenomenal orientation and perceived achromatic color 51
By William Epstein

Value orientations and preference for a minimax strategy 55
By Alvin Scodel

Forming and acting on impressions 63
By Celeste Mc Collough

(over)

Copyright, 1961, by The Journal Press
Provincetown, Massachusetts, U. S. A.
Entered as second-class matter January 26, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879
Second-class postage paid at Provincetown, Mass.

Task difficulty and anxiety level as factors affecting performance in a verbal learning situation 77
By Bernard W. Harleston and Scott M. Cunningham

The discontinuity in Schmeidler's ESP-Rorschach data 87
By R. A. Mc Connell

Organization of relevant and irrelevant words in the solution of verbal problems 99
By D. M. Johnson and E. R. Hall

The variation of choice reaction time among groups of hospitalized psychiatric patients 105
By Donald J. Dillon

Verbal conditioning: Discrimination and awareness 115
By Frederick H. Kanfer and John F. Mc Brearty

Trends toward multiple authorship in scientific publications 125
By Mildred B. Mitchell

How fat is fat? Some consequences of similarity between judge and judged object 133
By Samuel Fillenbaum

The sensory parallel of the reorganization period in the cortical response in intermittent retinal stimulation 137
By S. Howard Bartley, Thomas M. Nelson, and Jane E. Ranney

The geochronograph method for representing nonstationary electrograms, such as heartbeat and respiration 149
By Samuel J. M. England and Benjamin Pasamanick

Clinical perception of sexual deviates 157
By Leo H. Berman and Lawrence Zelic Freedman

The growth and extinction of expectancies in chance controlled and skilled tasks 161
By Julian B. Rotter, Shephard Liverant, and Douglas P. Crowne

The relationship between rational and irrational authoritarianism . . . 179
By Stanley A. Rudin

Delayed auditory feedback and ethanol: Effect on verbal and arithmetical performance 185
By Robert B. Forney and Francis W. Hughes

A modified version of the Wisconsin General Test Apparatus 193
By Allan M. Schrier

Meaningfulness (*m*) and transfer phenomena in serial verbal learning . . 201
By Clyde E. Noble

Toward cutaneous electro-pulse communication 211
By B. Von Haller Gilmer

Intractable asthma: Conflict of period of toilet training 223
By Harold A. Abramson

The influence of a distorted group norm upon judgments of adults and children 231
By Read D. Tuddenham

Measurements of simultaneous brightness contrast across the retina . . . 241
By William R. Mackavey, S. Howard Bartley, and Carmine Casella

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 3-24.

INJURY-PRODUCING PRIVATE MOTOR VEHICLE ACCIDENTS AMONG AIRMEN: PSYCHOLOGICAL MODELS OF ACCIDENT-GENERATING PROCESSES*[1]

Department of Psychology, The City College of New York and Public Service Research Institute, Stamford, Connecticut

JOSEPH E. BARMACK AND DONALD E. PAYNE

A. INTRODUCTION

An extensive literature on the causation and prevention of motor vehicle accidents has accumulated in recent years. A review by McFarland, Moore, & Warren (1955) cited 1031 articles. They are currently engaged in bringing the review up to date—the new review will cite more than 3000 articles (McFarland, 1960). Both the earlier review and the forthcoming one indicate clearly that, with very few exceptions, one after another theory or hypothesis about the general causation of accidents has been contradicted. Accident proneness, moderate sensory defects, deficiencies in simple psychomotor skills, low intelligence, etc., have each been suggested as general accident causes, but when carefully controlled studies have been conducted using adequate statistical procedures, the contribution of these hypothesized causes has turned out to be small or nonexistent. The exceptions, such as the contribution of alcohol, and heart failure, are relevant to only a limited portion of the accident causation picture.

There are a number of reasons why so many hypotheses have missed the mark.

1. The search for general causes is based on the assumption that accidents constitute a homogeneous "disease" with a specific etiology. There is a growing realization that this assumption is false.

2. One of the most important causes of research failure has been man himself. Man doesn't consciously want accidents.[2] He uses his resources

---

* Received in the Editorial Office on January 20, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported in part by the Office of the Surgeon General, Department of the Army, under sponsorship of the Commission on Accidental Trauma of the Armed Forces Epidemiological Board, and in part by research grant RG-7025 from the National Institutes of Health, U. S. Public Health Service.

The opinions expressed are the authors' and should not be construed as reflecting the views or endorsement of the sponsors or the military services.

[2] If an "accident" were planned in advance or deliberately and consciously brought about it would be considered vandalism (if only property damage resulted) or assault (if human injury resulted).

to avoid them. Other people on the road do the same thing. On uncommon occasions he fails. For methodological reasons, we have little information about the unique combination of psychological processes and acts which generate the failures.

Most accidents are first accidents. This fact limits the predictability of accidents from data on individual driver characteristics but not from other driving situation properties, e.g., characteristics of the road, the vehicle, the time of day, etc. (Uhlaner, 1960). If a driver is involved in two or more accidents, it is unlikely that the specific behavioral events in the first accident are identical with those in the ones that follow, although some of the intervening processes might be similar, e.g., emotional acting out, alcohol intoxication, etc.

3. In exploring the intervening processes, we may be demanding more advanced hypotheses from the contributing basic sciences than they are capable either of supporting or refuting. For example, the hypothesis of the causal contribution of an unconscious wish for an accident is exceedingly difficult to test. The crucial problem is not to demonstrate that the wish exists, but to show that a particular accident would not have occurred without it.

4. Even if we had the ability to be present at the time of the accident without our presence influencing the nature of the pre-accident processes, there probably would be wide variations of opinion on what to study, depending on the theoretical predilections of the investigator. This accounts, in part, for the variations in choice of factors to correlate with accidents. Many of the measures which have been studied have only fractional relevance to either successful or unsuccessful driving. Even driving simulator measures represent a small part of driving behavior of unknown importance in an environmental context of unknown relevance.

5. Investigators try to correlate failures in driving with successes on tests. People strive to make a test performance their best; the performances which preceded or coincided with the accident may be their worst.

6. Finally, accidents are grouped together to gain numbers in order to tease out common causes. We don't know which types of accidents can be grouped together to permit a meaningful causal analyses. Accidents can be grouped on the basis of time of occurrence, nature of the accident, consequences of the accident, properties of the driver, properties of the road, etc. They can also be grouped in terms of what the driver does immediately before the accident, such as falling asleep, travelling at high speed, passing, etc. Platt (1958) has presented a systematic enumeration of many of these factors. From what we already know, for each of these criteria for classifying

accidents, we are likely to get different distributions of correlated events and circumstances. That is to say, even though accidents are not consciously sought, the circumstances under which accidents occur are not merely distributed at random among all classes of accidents. Part of the problem of research failure lies in the fact that insufficient information is available on the extent to which distributions of accident-related properties vary as the accident classification changes.

Information of the above type on representative classes of accidents is needed. However, even if available, it would constitute only a partial solution. Accidents may be divided into those which are foreseeable and avoidable and those which are not, but there is overlap. Foreseeing and avoiding represent psychological processes preceding acts of omission or commission leading to the accident. We assume that these processes have more stability, i.e., that they are more likely to be repetitive than are the specific acts of omission or commission. Accordingly, the identification of these processes should contribute substantially to the detection of significant relationships between classes of accidents and to furthering our knowledge of the factors which contribute to them. How is such information to be obtained?

Lacking the ability to be present at the time of the accident, there appears to be no practical alternative to intensive post-accident investigation of the accident scene, of witnesses, and of surviving drivers. Methods of investigation of the physical characteristics of the post-accident scene have engendered confidence in their reliability. Photographic and measuring techniques and the experimental literature on vehicular behavior during crashes have contributed to this confidence.

In contrast, we mistrust the driver as a source of information. Impartial observers have demonstrated deficiencies in reporting high speed events reliably. There is doubt that the driver would be any more reliable. There is even greater doubt about the driver's ability to report inner processes. He is emotionally and defensively involved with police and insurance adjusters, among others. His self esteem as well as his head (and body) may have been seriously injured. He may suffer memory impairment from concussion, excessive alcohol, etc. In sum, he is an inelegant source of data. Little wonder that scientific investigators have generally avoided him, preferring to make assumptions about what went on before and during the accident. The fact remains, however, that for many inner processes, the accident driver is the sole source of information. In a study on the role of the federal government in highway safety, the following appears: "A great amount of useful information has, again, been established on human behavior in relation to influences and conditions. The only difficulty is that this knowledge has principally been gained in every type of surrounding except behind the wheel, in traffic. The influence on behavior, reaction, response, or ability, is known for many factors but their actual effect on the driver at the instant his action makes or breaks the collision course is not possible to ascertain in normal laboratory study" (U.S. Dept. Commerce, 1959, p. 36).

In addition to inner processes, for some behavioral data (for example, in a single-vehicle accident) the driver may be the only significant source of information. One challenge, then, is to make of the accident driver a valid source of data so that we may usefully tap his unique supply of information.

It has been evident for some time that in proportion to the total number of studies on automobile accidents, the number involving case studies designed to highlight the behavioral precursors and concomitants of the accidents is small. (E.g., Bletzacker, Baker, & Brittenham, 1957; Boek, 1958; Jones, 1958; Schmidt & Smart, 1959; Smith, 1938; Tilden, 1938.)

This paper has two objectives:

1. To describe the characteristics of an accident population which has sufficient homogeneity to be useful in identifying causal factors.

2. To consider theoretical explanations of the nature of the internal processes involved, using such empirical data as may have relevance to the theories.

In two previous reports (Barmack & Payne, 1961a; 1961b) the authors presented descriptive data on a representative sample of 138 airmen drivers involved in injury-producing accidents in privately owned automobiles. The data were obtained from lengthy interviews with the drivers and from accident report forms (AF Form 122) provided by Air Force accident investigators. Comparative data were obtained on 100 airmen controls selected to be representative of airmen drivers not involved in an automobile accident for at least one year.

Of the accident drivers, 64.5 per cent reported they had been drinking prior to the accident. The drinking was not an isolated event. The drinking accident group contained a larger proportion of frequent and problem drinkers and had a higher incidence of remote and recently disrupted home lives than the not-drinking accident drivers and the controls. However, there was so much overlap among the groups that no outstanding group differences emerged. This report presents additional analyses of group differences which have implications for explanation of accident causation.

B. RESULTS

1. *Choice of Homogeneous Accident Groups*

To obtain homogeneity the accident cases were limited, from the beginning, to those involving privately owned automobiles driven by airmen resulting in a lost-time non-fatal injury to the driver of the vehicle or in a fatal injury to the passenger. This class of accidents had the additional virtue that it was almost certain to be reported to the military authorities. Within this class of accidents we made further sub-divisions to secure specific high-homogeneity groupings.

A total of 77 of the 138 accidents were single-vehicle accidents. That is, they involved only the airman's vehicle (though eleven of the drivers alleged that the accident was preceded by an evasive action on their part to avoid a collision with another moving object). Therefore, the total sample was divided into two groups based on the type of accident: 77 single-vehicle accidents, and 61 multiple-vehicle accidents.

Preaccident drinking was reported by 76.6 per cent of the single-vehicle-accident drivers, and 49.2 per cent of the multiple-vehicle-accident drivers. The difference is statistically significant ($CR = 3.3$, $p = .001$). However, the 49.2 per cent does not include the incidence of drinking among the "other" drivers. Nevertheless, drinking appears to be more heavily involved in the single-vehicle than in the multiple-vehicle accident. This relationship is discussed in detail elsewhere. (Barmack & Payne, 1961a)

Accordingly we selected for analysis four groups of accident drivers, homogeneous with respect to type of accident and drinking: Fifty-nine single-vehicle-accident drivers who were drinking before the accident; eighteen single-vehicle-accident drivers who were not drinking before the accident; thirty multiple-vehicle-accident drivers who were drinking before the accident; and thirty-one multiple-vehicle-accident drivers who were not drinking before the accident.

Again, it should be noted that reference to drinking is limited to the airman driver, and does not include the driver of the other vehicle.

2. *Correlated Properties of the Four Classes of Accident Drivers*

Accident site data. To understand the processes preceding an accident requires knowledge of what the driver wanted to do. As a minimum, he wanted to stay on the road—with a few exceptions in cases in which the driver wanted to leave the roadway to avoid a collision with an oncoming vehicle.

Information about the site of the accident sheds some light on the conditions which may affect his success or failure. Table 1 shows some accident site characteristics associated with each of the four classes of accidents.

Three significant relationships are revealed by Table 1.

TABLE 1
DISTRIBUTION OF ACCIDENTS BY ACCIDENT SITE

	Vehicle Involvement			
Accident Site	Single-Vehicle (N = 77) %	Multiple-Vehicle (N = 61) %	Difference	p
Straightaway	35.1	47.6	—12.5	NS
Curve	51.9	8.2	+43.7	.0001
Intersection	13.0	44.2	—31.2	.0001
	100.0	100.0		
	Driver Condition			
	Drinking (N = 89) %	Not-Drinking (N = 49) %	Difference	p
Straightaway	33.7	53.1	—19.4	.03
Curve	41.6	16.3	+25.3	.003
Intersection	24.7	30.6	— 5.9	NS
	100.0	100.0		

1. Curves were the primary site for the single-vehicle accident.
2. Curves posed more of a problem for drinking drivers than for not-drinking drivers.
3. The straight road and the intersection were the most heavily represented sites of the multiple-vehicle accident, both combined accounted for approximately nine out of ten multiple-vehicle accidents.

The straight unintersected road appeared to be the major site for the not-drinking accident. We assume that, on the average, the unintersected straight road constitutes by far the largest proportion of road sites traversed by drivers.

The wide range of percentages of accidents on curves is particularly interesting, as shown in Figure 1. Curves were the accident site for from 3.2 per cent to 55.9 per cent of the accidents, depending upon certain characteristics of the driver and of the accident.

Why are curves so selective of single-vehicle accidents and of drinking drivers? Excessive speed might be one answer. Table 2 shows that travelling at excessive speed is more heavily represented in the drinking-driver, single-vehicle accident category than in any of the other three groups.

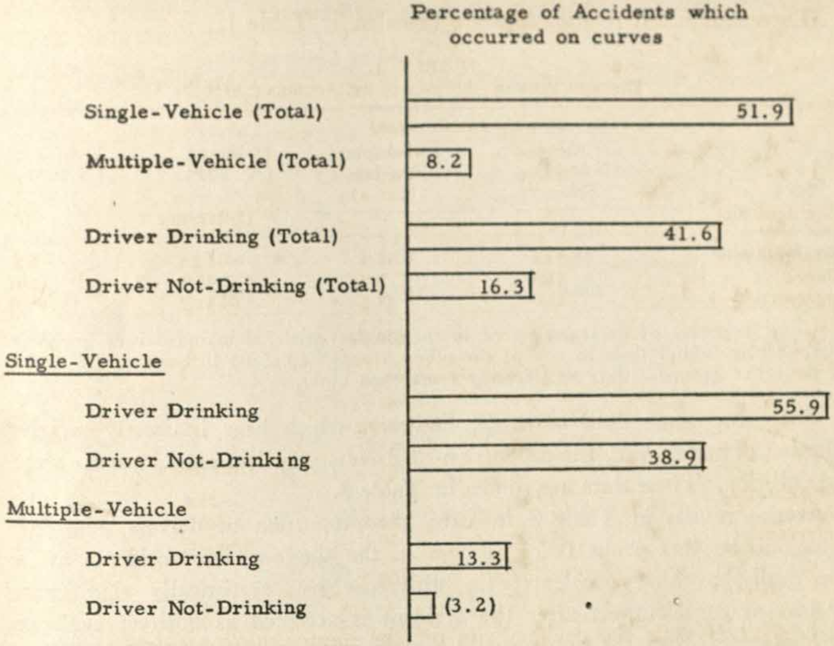


FIGURE 1
Percentages of accidents which occurred on curves for various classes of accidents, taking into account the nature of vehicular involvement in the accident and/or the preaccident condition of the driver.

Although Table 2 indicates an association between speeding, drinking drivers, and single-vehicle accidents, for those cases in which numerical estimates of speed were available, the median speed was only twelve mph higher for the drinking than the not-drinking group. In only four instances were the admitted speeds extraordinary, i.e., in the 90-120 mph range.

Speed data must be treated with caution. Official estimates of excessive speed in single-vehicle accidents are open to question. The very fact that an accident occurred is sometimes taken as *prima facie* evidence that the driver's speed was "excessive." Consequently, the accident investigator who reports excessive speed is unlikely to be challenged. On the other hand, evidence of drinking is more difficult to obtain and is more likely to be disputed (at least by the driver) in the absence of incontrovertible evidence.

Under these conditions investigators tend to use "excessive speed" as a catchall designation.

TABLE 2
REPORTED INCIDENCE OF SPEEDING

	Accident Classification			
	Single-Vehicle		Multiple-Vehicle	
Speed	Driver Drinking (N = 59) %	Driver Not Drinking (N = 18) %	Driver Drinking (N = 30) %	Driver Not Drinking (N = 31) %
Excessive	66.3*	38.9	33.3	25.8
Not Excessive	33.7	61.1	66.7	74.2
	100.0	100.0	100.0	100.0

* The incidence of excessive speed in the single-vehicle, drinking-driver group is significantly higher than in any of the other groups. ($p < .05$ in each case.) None of the other groups differs significantly from each other.

There are some data available, however, which bear indirectly on the question of speed, viz., the proportion of drivers who lost consciousness after the collision. These data are shown in Table 3.

As the results in Table 3 indicate, the proportion of drivers who lost consciousness was about twice as high in the single-vehicle accidents as in the multiple-vehicle accidents; the difference was statistically significant. If loss of consciousness after the accident is accepted as indirect evidence of high speed, then the results could be taken to support the conclusion that the single-vehicle-accident drivers were driving faster than multiple-vehicle-accident drivers. However, there are contaminating factors. Loss of consciousness may not be reported reliably. In addition, the proportion of drinking drivers who lost consciousness was higher than the proportion of not-drinking drivers who lost consciousness—and this difference was also significant. This association may be due to an inability to protect oneself in a crash, or may simply reflect a higher incidence of alcohol-induced "blackouts."

In spite of the ambiguity of the evidence on excessive speed in single-vehicle accidents, it seems evident that the drivers did not exert extra caution to compensate for possible effects from drinking. The interview protocols suggest that more important than speed, *per se,* was a failure to reduce speed on the curve. This failure turned up in 24 per cent of the interview protocols of the single-vehicle, drinking-driver accidents, and in 11 per cent of the single-vehicle, not-drinking ones.

The data presented so far show that, for the airman population, single-vehicle accidents have correlated site properties which are different from those of multiple-vehicle accidents. In addition to these *accident-site* differences, there are certain respects in which *single-vehicle-accident drivers* differ from *multiple-vehicle-accident drivers.*

TABLE 3
LOSS OF CONSCIOUSNESS FROM COLLISION

	Vehicle Involvement			
Item	Single Vehicle (N = 77) %	Multiple Vehicle (N = 61) %	Difference	p
Lost consciousness after collision	63.6	32.8	+30.8	.0003
Did not lose consciousness after collision	36.4	67.2		
	100.0	100.0		
	Driver Condition			
	Drinking (N = 89) %	Not-Drinking (N = 49) %	Difference	p
Lost consciousness after collision	56.2	38.8	+17.4	.05
Did not lose consciousness after collision	43.8	61.2		
	100.0	100.0		

Marital status. We have reported previously (Barmack & Payne, 1961b) that the drinking-driver accident sample was overrepresented with airmen who were from broken homes or who were living apart from their wives. The not-drinking accident sample was, if anything, underrepresented in this respect (in comparison with a control group).

In a subsequent analysis in which the type of accident (single- versus multiple-vehicle) was used to fractionate the data, some additional relationships became apparent. These data are presented in Table 4.

For this sample the single-vehicle accident seems to be primarily a single airman's ailment, and the multiple-vehicle accident seems to be a married airman's ailment. This rather odd coincidence may be understandable in terms of the sociology of exposure of the two groups. There is likely to be a higher proportion of preaccident drinking, including bar and party hopping, among the unmarried as compared with married airmen. The married airman's driving may be limited to residential, urban areas, whereas the unmarried airman's driving may cover interurban as well as urban areas.

TABLE 4
MARITAL STATUS OF DRIVER AND TYPE OF ACCIDENT

	Vehicular Involvement	
Marital Status	Single-Vehicle (N = 77) %	Multiple-Vehicle (N = 61) %
Married	40.2	65.6
Single	59.8	34.4
	100.0	100.0

Note.—$X^2 = 7.74$, df = 1, p = .005. This value of X^2 supports the hypothesis that there is an association between marital status and type of accident.

Family background. The relationship between family background and type of accident is presented in Table 5.

Table 5 shows that the highest incidence of home disruption occurred

TABLE 5
HOME DISRUPTION AND TYPE OF ACCIDENT

	Single Vehicle		Multiple Vehicle	
Group	Drinking (N = 59) %	Not-Drinking (N = 18) %	Drinking (N = 30) %	Not-Drinking (N = 31) %
Home Disrupted	47.4	27.7	66.7	29.0
Home Not Disrupted	52.6	72.3	33.3	71.0
Total	100.0	100.0	100.0	100.0

Note.—The "home disrupted" includes four categories: (1) unmarried who come from a broken home, (2) married and living with wife but comes from a broken home, (3) married not living with wife and comes from a broken home, (4) married not living with wife and does not come from a broken home. The "home not disrupted" includes two categories: (1) single and does not come from a broken home, (2) married living with wife and does not come from a broken home.

A X^2 test applied to the drinking and not-drinking groups (i.e., ignoring type of accident) resulted in the following: $X^2 = 7.22$, df = 1, p = .01. This value of X^2 supports the hypothesis that there is an association between family background and preaccident drinking.

among airmen in the drinking-driver, multiple-vehicle accident. Both drinking-driver groups are clearly different from the not-drinking ones in incidence of home disruption. However, the proportion of home disruption in the drinking-driver, single-vehicle accident category is proportionately small because of a high concentration of unmarried individuals from non-disrupted homes.

Table 6 compares the proportions of unmarried airmen from non-disrupted homes among the four accident categories.

TABLE 6
UNMARRIED AIRMEN FROM NON-DISRUPTED HOMES

	Single Vehicle		Multiple Vehicle	
Group	Drinking (N = 59) %	Not-Drinking (N = 18) %	Drinking (N = 30) %	Not-Drinking (N = 31) %
Unmarried airmen from non-disrupted homes	40.7	55.5	13.3	25.8
All others	59.3	44.5	86.7	74.2
	100.0	100.0	100.0	100.0

Note.—The proportion of unmarried airmen from non-disrupted homes among the drinking drivers involved in multiple-vehicle accidents was significantly lower than among the single-vehicle-accident drivers, whether drinking (CR = 2.6, p = .01) or not-drinking (CR = 3.1, p = .003). The proportion of unmarried airmen from non-disrupted homes among the not-drinking drivers involved in multiple-vehicle accidents was significantly lower than among the not-drinking drivers involved in single-vehicle accidents (CR = 2.1, p = .03). None of the other differences was significant.

The drinking driver involved in a single-vehicle accident appears to be drawn from either one of two different populations:

1. The home disrupted population.
2. The home non-disrupted, but unmarried population.

The role of drinking in the two groups may be quite different. In the home-disrupted group, drinking may serve as a tranquilizer—a relief from psychic pain generated by the disruption. For the unmarried group the drinking may serve the purposes of reality testing, demonstration of masculinity, socializing, etc. These hypotheses warrant investigation in future research.

In the foregoing we have shown that when our sample of accident drivers is grouped according to whether they were involved in single-vehicle or multiple-vehicle accidents and whether the accident was or was not preceded by drinking, greater homogeneity is achieved in a number of correlated properties. In the following section we will consider some hypotheses about intervening processes in these categories of accidents.

3. *Preaccident Processes*

Active versus passive psychological processes. The drinking-driver, single-vehicle accident, associated with speeding, disrupted homes, unmarried airmen, etc., may suggest accident generating processes in the nature of *active*, symbolic, self-destructive acts released by alcohol. Our data, however, do not support this hypothesis. Rather they suggest that most (indeed not all) of these accidents were an end product of passive, not active psychological processes. Furthermore, the intent was not necessarily self-destructive. The evidence for this position follows:

Impaired consciousness. A content analysis was made of the interview protocols to identify perceptual, cognitive, motor, and emotional factors which appeared in the description of the accident events.

Those factors which reflect impaired or maladaptive characteristics of consciousness are shown in Table 7.

TABLE 7
IMPAIRED OR MALADAPTIVE CHARACTERISTICS OF CONSCIOUSNESS AMONG ACCIDENT DRIVERS

	Accident Classification			
	Single-Vehicle		Multiple-Vehicle	
Impairment	Driver Drinking (N = 59) %	Driver Not Drinking (N = 18) %	Driver Drinking (N = 30) %	Driver Not Drinking (N = 31) %
Perceptual surprise	35.6	16.6	33.3	16.1
Fell asleep or passed out before collision	35.6	38.8	6.7	9.7
Memory gap for pre-accident events	18.6	11.1	6.7	9.7

A broad range of impairment is noted, from "perceptual surprises" through "memory gap for preaccident events," where there is no recall of events well before the accident happened. Some of these categories are tallied for each person more than once when appropriate, consequently the percentages do not necessarily add up to 100 per cent. For the total single-vehicle group (disregarding preaccident drinking or not-drinking) there is a high percentage, 36.4 per cent, of reports of falling asleep or passing out before the accident. For the total multiple-vehicle group, only 8.2 per cent reported falling asleep or passing out before the accident. The difference between the groups is highly significant ($p = .0001$).

Memory deficit may be assumed to involve either more serious impairment at the time of the accident, or reporting contaminated by retrograde amnesia. An example of a report classified as a memory deficit is: "I started to go to the bar to pick up my buddy. I remember I stopped for a red light. I don't remember anything after that."

One of the interesting findings in Table 7 is the higher incidence of perceptual surprises among the drinking as compared with not-drinking accident drivers. Examples of content classified in this category are:[3]

"I was driving along the highway at normal speed—55 mph. I was sleepy but by no means asleep. *All of a sudden* a car loomed up in front of me and I hit it."

"I ran straight into the caution light. I should have turned right. *I didn't see the blinker light until the last minute.* I was doing about 60. I normally do about 35 around here. My buddies didn't scream or warn me."

"I was doing around 50-60 when I came up to highway 76. I knew there was a stop sign. I was looking to the left *but it came up so quick!* I applied my brakes and cut my wheel to the left at the T intersection. The car just kept going. I seen I was going to the tree and that was it."

One is surprised by events which develop at a pace that exceeds normal expectancy. Events may develop too quickly because one is travelling at excessive speed, because an opposing vehicle comes on the scene suddenly, or because of distraction.

Degraded anticipation. It has been shown (Table 2) that the drinking driver failed to be more cautious than the not-drinking one. He was likely to travel at somewhat higher speeds or to fail to slow down for a curve. There appeared to be no significant differences among all groups of drivers in frequency of reported evasive actions. In addition, only 3 per cent of the total of 138 accident drivers reported having been distracted. Accordingly, a degradation of anticipation—in relation to the pace set by the driver—must carry the main burden of explaining the higher incidence of perceptual surprises among drinking drivers. How does the degradation come about?

We assume that safe driving requires *active* searching for cues to danger, the development of hypotheses about likely hazards, and motor adjustments to these signs and estimates. The hypotheses generate sets which selectively sensitize the driver to relevant information about the self, the car, and the environment. These activities go on at high speed, semiautomatically, in a diffuse informal way rather than as precise formal intellectual processes. F. Allport (1955) provides a detailed review of these processes, though not in relation to driving.

---

[3] Italics added in the quoted reports.

To some extent, one may be able to function safely at slow speeds in spite of a *passive* level of functioning of these processes. Alcohol apparently induces a passive level of functioning without creating, at least in some drivers, the intention to reduce speed. However, higher speeds require more active anticipatory processes. The lack of adjustment to this higher demand level or the inability to sustain more active processes, defeated most of the drinking drivers.

There is a broad range of performance capability between peak alertness and deep sleep levels (Kleitman, 1957; Malmo, 1959). This range must be accompanied by different levels of mental organization in which emotion, sensory stimulation, drive, interest, and effort all play somewhat different and overlapping roles. The term "activation" has been suggested by Lindsley (1951) to describe phenomena which do not guide but simply "energize" activity. Vigilance describes some optimum level of this state associated with low error performance. Boredom (Barmack, 1937), mental blocking (Bills, 1931), certain types of satiation and skill fatigue (Bartlett, 1943) refer to low levels of these activities in relation to the task at hand.

The internal state most clearly associated with the accidents in our cases was a downward shift (i.e., toward sleep) in the level of activity along the sleep-vigilance continuum. Support for this view comes from the fact that out of 39 drivers who had single-vehicle accidents on turns, 64.1 per cent did not react to the curve at all, but went off the road in a straight line without evidence of fighting it until they were well off the road. Another 20.5 per cent reported they actively struggled with the car's direction but were unsuccessful.

Flagging vigilance among single-vehicle-accident drivers is suggested by still another trend. For those cases on whom there was specific information, a left hand turn was involved in more lost-time single-vehicle accidents than a right hand turn. The number of cases is small and not statistically significant, i.e., nineteen had their accidents on left hand turns, and fifteen on right hand turns. However, we suggest as a hypothesis that the geometry of the right hand turn permits a longer safe passage in which to assess course error than a left hand turn.[4]

The concept of passivity of central processes also may be inferred from some data from multiple-vehicle accidents. It is reflected in different proportions of "who ran into whom" among the drinking and not-drinking accident drivers. The following criteria were used to identify the "rammer": (*a*) The front end of his car hit. (*b*) In the case of a head-on collision, his car was in the wrong lane. (*c*) In a sideswipe his car was in motion or was passing.

The type of impact is apparently a sensitive, as well as an objective, index of drinking. Table 8 shows that the drinking driver was more likely to do the ramming than the not-drinking one.

TABLE 8
"WHO RAN INTO WHOM?" IN MULTIPLE-VEHICLE ACCIDENTS

	Driver Condition			
Group	Drinking (N = 30) %	Not-Drinking (N = 31) %	Difference	p
Airman rammed other vehicle	80.0	51.6	+28.4	.02
Other vehicle rammed Airman	20.0	48.4		
	100.0	100.0		

The not-drinking driver had an equal probability of hitting or being hit. The drinking driver ran into others four times as often as he was run into. The proportion of rammers among the drinking drivers was significantly greater than the proportion among the not-drinking drivers at the .02 level.

Among the 61 multiple-vehicle accidents, there were sixteen head-on collisions. Table 9 shows the distribution of lane positions as a function of drinking for this group of sixteen drivers.

The number of cases involved is small, but the trend is clear. The drinking driver was preponderantly in the wrong lane. Incidentally, the single drinking driver who was in his own lane was hit head-on by a drunken driver who had crossed over.[5]

---

[4] This question of curve geometry is not ordinarily taken into account in published collections of accident statistics. For instance, the National Safety Council's annual *Accident facts* (1960) includes tabulations of accidents associated with "improper turns" (p. 52) and "level road" versus "hillcrest" or "on grade" (p. 57), but does not include breakdowns for accidents on curves.

[5] The fact that almost all the drinking drivers had crossed into the path of

The typical preaccident conditions of these classes of accidents under discussion seems to have little to do with organic sensory defects, lack of driving skill, poor intellectual assessment of the risks in critical road choice situations, or some self-destructive act released by alcohol. Excessive speed is a frequently reported correlate of the single-vehicle accident, but most of the "speeding" may be appropriate for the road conditions (if not for the individual). Speeding suggests activity and its presence may lead us to overlook the existence of a more important psychological process—passivity of those sets necessary for sustained vigilance.

TABLE 9
LANE IN WHICH AIRMAN WAS RIDING IN MULTIPLE VEHICLE ACCIDENTS

Group	Number of Cases Drinking (N = 10)	 Not-Drinking (N = 6)
Airman riding in wrong lane	9 (90%)	2 (33%)
Airman riding in correct lane	1 (10%)	4 (67%)

Note.—Because of the very small number of cases, the comparison of proportions is an inappropriate statistic (McNemar, 1949). However, the Fisher Exact Probability Test (Siegel, 1956) is applicable. When applied to these data the Fisher test results in $p = .04$.

Unconscious directing processes. The model presented so far has been concerned primarily with energic or supporting rather than directing processes, but is there any evidence in the interview data that unconscious directing processes may have generated some of the accidents? Such evidence emerged in many cases. But physical trauma as an *aim* of these processes turned up rarely. The rarity may be a function of the limited time spent interviewing the driver. Unconscious aims did appear as contributors, but their directions were varied, and the accidents appeared incidental to them. A relatively clear relation between the accident and wished-for physical trauma may be inferred from the following protocols.

In the first illustration the airman left his pregnant wife at a party in order to take another woman for a ride in his yellow convertible:

> "We were coming back to the party. I came to this intersection three blocks from the party. I hadn't decided yet whether to go back or to cruise around and try to make her. I felt like a horse's butt leaving

---

oncoming traffic could be interpreted as either a consequence of passivity, or an unconsciously directed act of a self-destructive or other-destructive nature. It is impossible to decide between these hypotheses on the basis of the results presented in Table 9. However, we believe the passivity-hypothesis is more consistent with the weight of evidence presented throughout this article for most of the cases.

> my wife. I kind of hated her too for getting me excited. I approached the corner at about 50 mph. I don't know if I thought I saw lights or I thought a car might come and I jammed on the brakes. She slipped off and hit her cheek on the dashboard. There was no car. We went back to the party."

One might infer that the accident was a hostile act which resolved a conflict for the airman.

In the following case there is evidence of emotional disturbance but the accident appears to be the outcome not of a specific desire for it but rather that the turbulence found expression in driving behavior which served as a tension releasing device. The accident occurred because the form of tension release was maladaptive.

The driver spent all day preparing an apartment for his fiancee who was flying in from abroad. He described himself as being very excited, jittery and nervous. He drank to calm down. He was joined by a buddy. He delayed going to the airport and then left in a panic that he would be late:

> "We were doing about 90. ————asked me to slow down. I came to the intersection and saw two cars coming from opposite directions. I didn't see them until I was at the intersection. I thought they had their brights on. I wished the fool on my right would stop, but I knew he wouldn't. I veered to the left and hit the traffic island and passed out."

Another case which reflects tension release as a primary factor is the following. This airman was keeping company with a woman separated from another airman. The night of the accident she told him she didn't feel well. They argued and he left "to go out with the boys." Late in the evening he decided to see her. When he drove past her house he saw her estranged husband's car parked in front of the house:

> "This made me mad. She was supposed to be separated and she said she was sick. I drove past. I hated women. I stepped on the gas. I must have been doing 100 mph. A big Chrysler was making a turn and his tail was sticking out. I got maybe 100 feet away so I swerved to the right. I hit the curb and lost the wheel. The car headed over the island and ran into a palm tree."

Another example of tension release is reflected in the following:

> "I was going about 55 to 60 and saw a car down the road that wasn't going fast enough to suit me. It looked like he'd stopped. I said to myself, 'that son of a bitch has stopped. I'll pass him.' I swung out and started to pass but I saw a car coming so I came back and slammed on the brakes. I tried to hit him square so I wouldn't roll."

A number of points are illustrated by this narrative. First, his driving apparently was guided by a generalized psychological tempo or "press" which he seemed to be reluctant to modify. Second, the decision to pass was not made merely on the basis of deliberation or assessment, but rather contained a large component of impulsive and aggressively-tinged behavior.

The concept of "press," for instance is illustrated also in the following drinking single-vehicle accident. The driver had been drinking most of the afternoon and early evening with some friends. On his way home:

> "I felt guilty about not letting my wife know. I was pushing, doing 50-55 when I thought I could in that snow. I don't think I was doing more than 40 at that particular part of the road. I started to fishtail and made the mistake of stepping on the brake. The car spun around."

Still another example of tension release is reflected in the following case:

> "I was at this party. I had a date with a girl at 12 (midnight). I was in a hurry. I went around the house, made a peel out turn at the corner. I started to speed up in ————. I passed the red light. The road narrowed and I hit the car parked on the shoulder. When I hit I passed out. I think I was doing 55-60. The police say 65-70."

There are other cases where the accident is accompanied by more complex acting out processes in which the accident may or may not be part of the aim of the sets or attitudes. In one, the airman provoked a chase by an air police patrol at night. He attempted to elude it by putting his lights out and turning into a side street where he ran into a concrete stanchion he did not or could not see. This chase-pattern was one of his pre-service hobbies with local sheriffs. The accident was incidental to the pattern, but the pattern itself increased his vulnerability.

Another illustration is that of a father of two children who had been involved in three drinking accidents, but had been hurt only in the last one. He was the middle of thirteen children born to a blind mother and a father who would go on sadistic drinking bouts two or three times a month:

> "You either had to kill him or get out. One time I got so mad I took a gun to him. My younger brother hid him out at a neighbor's for three days. When he came back he swore off and he has kept his pledge. I left."

On the night of the last accident he ran into an embankment at 50 mph in a 25 mph zone while trying to avoid a post. (The interviewer and the airman returned to the scene. The post proved not to be there.) His drinking may have been the result of an identification with his hated father (A. Freud, 1937; Mowrer, 1950). If this pattern did reflect defensive identification with a sado-masochistic father, the airman's accidents might be viewed as masochistically inspired acts, as compared with the father's sadistic pattern (Fromm, 1941).

There is a strong correspondence between the experiences narrated by the preceding drivers and Tillmann's (1960) analysis of the use of the automobile as "a narcissistic extension of the self." On the basis of his experiences with accident repeaters required to attend group psychotherapy sessions under threat of suspension of their driving licenses, Tillmann concludes:

> "It is noticed that the group members from time to time have recounted a feeling of rage while driving their automobiles, especially following a situation where they have felt a loss of identity. They indicate a need to express themselves through a medium where they feel they have control over some portion of their feelings. In these instances the automobile has now become an extension of themselves, and they drive it in such a fashion as to channel out these feelings of rage, through a medium where they feel they have mastery. They also fuse themselves with this falsely-acquired power and attempt to exercise mastery over their driving environment as it best suits their needs, *rather than the realistic needs of the driving environment*" (Tillmann, 1960, pp. 4-5).[6]

In addition to the aggressiveness noted by Tillmann, we would include other generalized tension release mechanisms.

In the preceding discussion we have examined such data as were available to suggest inferences about preaccident processes. We inferred they were predominantly passive rather than active and that the unconscious aim of the preaccident behavior was more often tension-reducing than self-aggressive.

D. Rapaport (1960) suggests that although some might argue that by using lengthier psychoanalytic interviews underlying "self-destructive" impulses may be revealed, this should not be expected in all cases, even theoretically. Some of the impulses will be manifestly self-destructive, some of them will not be reducible to such, and some of them will be of an entirely different nature. He raises another possibility—that the reduction in ego autonomy (or surrender of active processes)[7] may in itself be an instrument of self destruction. But this conception of self destructiveness would be entirely different from those which have been used in the past. In any event, such a possibility further complicates the task of establishing the extent to which an unconscious self-aggressive aim is essential for the occurrence of an accident.

---

[6] Italics added.

[7] Ego autonomy might be more completely described as the ability to engage in chosen goal-directed behavior without perceptual, cognitive, or motor compulsion. Active sets are an essential condition for this ability.

C. SUMMARY

Correlations between postulated accident causes and accident experience have been disappointingly low. A number of factors to account for these failures of prediction were discussed and two of them were singled out for special attention.

1. The tendency to treat accidents as if they were homogeneous.
2. The paucity of studies concerned with both behavioral and intervening processes at the time of the accident.

The subjects of this study were 138 airmen drivers involved in automobile accidents resulting in lost-time injury. Roughly half of the accidents studied were single-vehicle accidents; this class of accident exhibited otherwise different correlated properties than the multiple-vehicle accidents.

The single-vehicle accidents had a substantially higher incidence of drinking drivers, loss of driver consciousness after impact, and reported speeding. The speeding may not have been inappropriate to the road characteristics. Curves were the primary location of this class of accidents, whereas straight roads and intersections were the primary sites of the multiple-vehicle accidents.

While the drinking-driver, multiple-vehicle accident group drew heavily from a population with a home-disrupted background, the drinking-driver, single-vehicle accident group drew equally from two populations of drivers —those who come from disrupted homes and those who are single but do not come from disrupted homes. Drinking probably serves different functions for the two groups. We assume that it serves as a tranquilizer for the home disrupted and serves other needs for the single individuals not from disrupted homes.

Evidence is presented to support the view that the primary preaccident psychic processes for the classes of accidents studied are passive rather than active. While evidence for the operation of a variety of unconscious processes was obtained, the interview protocols do not support the view that the aim of these processes was necessarily to achieve a self-aggressive, destructive end. Rather in many instances their aim appeared to be tension discharge inappropriate to the reality requirements of the road. However, the possibility remains that the shift from an active to passive psychic processes may be itself an expression of a destructive aim.

REFERENCES

1. ALLPORT, F. H. Theories of Perception and the Concept of Structure. New York: Wiley, 1955.
2. BARMACK, J. E. Boredom and other factors in the physiology of mental effort. *Arch. Psychol.*, 1937, No. 218, 67-74.
3. BARMACK, J. E., & PAYNE, D. E. Injury-producing private motor vehicle accidents among airmen: I. The role of drinking. *Hwy Res. Bd. Bull.*, 1961a, in press.
4. ———. Injury-producing private motor vehicle accidents among airmen: II. Background correlates of the lost-time accident. *Hwy Res. Bd. Bull.*, 1961b, in press.
5. BARTLETT, F. C. Fatigue following highly skilled work. *Proc. Roy. Soc., Ser. B.*, 1943, **131**, 247-257.
6. BILLS, A. G. Blocking: A new principle of mental fatigue. *J. Psychol.*, 1931, **43**, 230-245.
7. BLETZACKER, R. W., BAKER, R. F., & BRITTENHAM, T. G. A study of one-car accidents. Columbus, O.: Ohio State Univ. Engng. Exp. Stat., 1957.
8. BOEK, J. K. Automobile accidents and driver behavior. *Traff. Safety Res. Rev.*, 1958, **2** (9), 2-12.
9. FREUD, A. The Ego and the Mechanisms of Defense. London: Hogarth, 1937.
10. FROMM, E. Escape from Freedom. New York: Farrar & Rhinehart, 1941.
11. JONES, E. R. Planning for intensive investigation of behavioral aspects of traffic accidents. Evanston, Ill.: Northwestern Univ. Traff. Inst., 1958.
12. KLEITMAN, N. Sleep, wakefulness, and consciousness. *Psychol. Bull.*, 1957, **54**, 354-359.
13. LINDSLEY, D. B. Emotion. In S. S. Stevens (*Ed.*), *Handbook of Experimental Psychology*. New York: Wiley, 1951.
14. MCFARLAND, R. A. Personal communication, March 17, 1960.
15. MCFARLAND, R. A., MOORE, R. C., & WARREN, A. B. Human variables in motor vehicle accidents. Boston: Harvard Sch. Publ. Hlth., 1955.
16. MCNEMAR, Q. Psychological Statistics. New York: Wiley, 1949.
17. MALMO, R. B. Activation: A neuropsychological dimension. *Psychol. Rev.*, 1959, **66**, 367-368.
18. MOWRER, O. H. Learning Theory and Personality Dynamics. New York: Ronald, 1950.
19. NATIONAL SAFETY COUNCIL, STATISTICS DIVISION. Accident Facts: 1960 Edition. Chicago: National Safety Council, 1960.
20. PLATT, F. N. Operations analysis of traffic safety: A classification of traffic situations. *Traff. Safety Res. Rev.*, 1958, **2** (4), 17-24.
21. RAPAPORT, D. Personal communication, July, 1960.
22. SCHMIDT, W. S., & SMART, R. G. Alcoholics, drinking and traffic accidents. *Quart. J. Stud. Alcohol.*, 1959, **20**, 631-644.
23. SIEGEL, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.
24. SMITH, W. S. Skilled investigation at the scene of the accident needed to develop causes. In *Motor vehicle traffic conditions in the United States. Part 2.* Washington, D. C.: Govt. Printing Off., 1938.

25. Tilden, C. J. Case histories of fatal highway accidents. Washington, D. C.: Govt. Printing Off., 1938.

26. Tillmann, W. A. Group therapy amongst people involved in frequent accidents. Report to the Commission on Accidental Trauma of the Armed Forces Epidemiological Board, 1960.

27. Uhlaner, J. E. Personal communication, July, 1960.

28. U. S. Department of Commerce. The federal role in highway safety. Washington, D. C.: Govt. Printing Off., 1959.

Public Service Research Institute, Inc.
91 Prospect Street
Stamford, Connecticut

Published as a separate and in *The Journal of Psychology*, 1961, 52, 25-45.

PERCEIVED RATE OF ELECTRICAL PHOSPHENES AS DETERMINED BY MATCHING SQUARE WAVE AND SINE WAVE FLICKER RATES*[1]

The Training School at Vineland, New Jersey

JOHS. CLAUSEN AND RATHE KARRER

In a previous study by the senior author (Clausen 1955) a preliminary attempt was made to investigate the relationship between frequency of the stimulating sine wave current and the rate of the perceived flicker. This study seemed to indicate the following two points: 1. At stimulus frequencies ranging from 30 to 60 cps there is no direct relationship between the frequency of the stimulus and the resulting flicker rate. The flicker rate produced by the various frequencies are very poorly distinguished. There is, however, a direct relationship for stimulus frequencies below 10 cps. 2. At stimulus frequencies of 5 cps and below it is possible to count the number of flashes which result from the sine wave stimulation. There appears to be a close relationship in this frequency range, i.e., 5 cps sine wave stimulation produced 3.5 flashes per second. Further investigation made us believe that in the low frequency range there is a 1 : 1 relationship between stimulus frequency and perceived flicker rate, but that it is difficult to count accurately even at rates of 5 per second.

The present study investigates the relationship between frequency of square wave[2] stimulation and the resulting rate of perceived flicker. Since a full sine wave cycle, including a positive and a negative phase, gives a single flash, the question arises whether a comparable square wave cycle also acts in this manner. This is all the more interesting because we know that monophasic square waves of either polarity produce phosphenes. A further problem is to explore the factors which influence flicker rate. We have attacked this problem by varying square wave characteristics and determining the resultant flicker rate.

---

* Received in the Editorial Office on January 20, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was supported by National Science Foundation Grant G-2106 and G-7029.

[2] Square wave is here used as a generic term denoting a wave form having an essentially zero rise and decay times with constant amplitude values, rather than in its specific sense denoting a wave form having two constant amplitude values of equal duration.

A direct way of determining the rate of subjective flicker is not available beyond the point where the flashes can be counted. An indirect avenue seems to be in comparing sine wave stimulation to square wave stimulation, where pulse interval, pulse duration and polarity may be controlled independently.

A. Apparatus and Procedure

The square wave generator utilized was built according to specification by Barger and Gebhard (1952). In addition to independent control of amplitude, pulse interval and pulse duration, this generator produced the following four waveforms: 1. Positive pulses; 2. Negative pulses; 3. Alternating pulses with equal interval between pulses; and 4. Biphasic pulses, with zero interval between positive and negative pulses, so that repetition rate is controlled by the interval between the negative and positive pulses. The source of the sine wave stimulus was a Hewlett-Packard Low Frequency Oscillator (Model 202 C). The electrical stimuli were constantly monitored by a Hewlett-Packard Electronic Counter (Model 522 B) and a Dumont Oscilloscope (Model 304-A). The *S* was equipped with a three way selector switch allowing him to choose between sine waves, square waves and no stimulus.

Two silver electrodes, 25 mm. in diameter were placed on the right temple and the center of the forehead. Contact was assured by cleaning these areas with acetone before applying bentonite electrode paste. The electrodes were secured with scotch tape.

The procedure employed was the psychophysical method of Constant Stimuli with three categories of response. The square wave stimulus was always used as a reference, i.e., kept constant during a particular series of comparisons, since this stimulus was considerably more complicated to calibrate than the sine wave stimulus. In comparing two stimuli *S* was free to go back and forth between the two sources as much as he wanted before giving his judgment.. The square wave stimulus was compared to a series of sine wave stimuli 1 cps apart, which were presented in random order. The *S* judged whether the flicker resulting from the sine wave was slower, equal to or faster than that resulting from the square waves. A series was complete when three adjacent frequencies at the upper end of the series elicited judgments of faster and three adjacent frequencies at the lower end elicited judgments of slower. The number of stimuli in each series ranged from six to 28 with six to nine as the usual value. The upper extreme resulted from difficulty in judgment at the two highest frequencies, 25 and 30 cps. Occasionally a long series resulted from an inadequately chosen starting frequency. The *S* controlled the amplitude of the square waves himself, and requested the examiner to increase or decrease the amplitude of the sine waves for the purpose of matching intensity for each judgment.

Because of the impressive results of Gebhard *et al.* (1956) in determining difference thresholds for flicker rates of electrical phosphenes, we tried their procedure, the psychophysical method of Adjustment. Our experience with this method, however, was that it gave highly variable results, for which reason we preferred the Constant Stimuli method. Whereas the method of Constant Stimuli may not be adequate to determine Difference Limens, it seems well suited for determining Point of Subjective Equality (PSE) which measure was our primary concern. In the great number of matches performed it has happened only a few times that a subject made identical judgments for four repetitions. Thus it appears that steps of 1 cps in the variable stimulus were adequate.

The two authors served as *S*s. Before the present study both had considerable experience at introspective observation of electrical phosphenes, and during the course of the study became rather experienced in observing flicker rates. Both *S*s served for all the experiments. Any one match was repeated four times for each *S* so that each point of the following curves includes eight determinations. The four repetitions were consistently done on four consecutive days.

PSE was computed according to formulas cited by Woodworth (1938, p. 419) and the mean for the two *S*s determined. It is such means which have been plotted in the graphs of the Results section. The reported PSEs varied generally over the four repetitions from 0 to ± 1.5 cps, except for the highest two frequencies, 25 and 30 cps, which occasionally varied as much as ± 3.0 cps.

Unless otherwise specified matches were obtained for two light conditions, dark adapted and .01 mL reflected by the wall which the *S* faced.

It is no easy task to match the frequency of the flickering phosphenes. In the beginning of the study the difficulty in concentrating on the flicker rate became quite frustrating. With more experience this difficulty disappeared completely. But there are other factors which probably interfered with matching precision. Under certain conditions, such as low frequencies and short pulse duration of the square wave, there is a definite quality difference in the flicker from the two sources. Sine wave stimulation tends to be perceived as pulsating light intensity while the square wave stimulation is perceived as trains of discrete light flashes. There are, however, conditions where the phosphenes from the two sources are practically indistinguishable, e.g., frequencies around 20 cps, when the pulse duration of the square waves is around 20 ms.

A further difficulty stems from the fact that in many of our experiments we selected 4 ms duration for the square waves. This allowed us to include higher repetition rates than would have been possible with a longer pulse duration. We also wanted to have the pulse duration short in order to prevent the possibility of a double flash for a single stimulus, i.e., break as well as make-effect. With a rather low duty cycle, however, the phosphenes are not easily seen, and this has to be compensated for with increased intensity. This results in marked cutaneous sensations, and the rhythmic pounding on the head under the electrodes is a distraction when one is trying to determine the rate of the flicker.

Also, the rate of the phosphenes is somewhat different in central and peripheral parts of the field, being generally slower in the periphery. When the phosphenes are not very pronounced, it is a temptation to search in the field to find where they are most noticeable, and thus be somewhat inconsistent with regard to location where the match is made. In general, however, the observations were made in the upper quadrants of the field, about 45° off center. At times a square wave stimulation seems to give rise to two flicker rates. One of the two flicker rates seems more pronounced than the other, but it is not always the same flicker rate that is dominating on different days. Again, it is in the low frequency range where this interference is most likely to occur.

In spite of all these factors reducing the precision and stability of our matchings, we think we have a sufficient amount of data with sufficient consistency that significant trends are revealed.

In the course of our study a number of separate experiments were performed as one experiment pointed out the necessity of another. In the Results section below, we have, in general, followed the chronological sequence of the experiments.

The various problems which have been considered in the study are as follows: 1. Stimulus Polarity. Positive and negative square waves matched to sine waves. 2. Varying frequency in terms of cps. 3. Varying pulse duration at fixed frequency. 4. Varying frequency in terms of pps. 5. Varying light adaptation state of the eye during the matching procedure.

B. RESULTS

1. *Stimulus Polarity*

As already mentioned we considered four different types of square waves: positive pulses, negative pulses, alternating positive and negative pulses and biphasic pulses. In order to determine whether stimulus polarity influenced flicker rate a comparison was made between positive pulses and negative pulses for frequencies of 5, 10 and 20 pps at a pulse duration of 20 ms, utilizing both standard light conditions, i.e., dark adaptation and adaptation to .01 mL. The comparison was in terms of the sine wave frequency matched to the positive and negative square waves. The results are presented in Table 1. This table shows no systematic difference between positive and negative pulses. For the sake of economy we therefore excluded the negative pulses from subsequent experiments under the supposition that the two different pulses would not show different results.

TABLE 1
SINE WAVE FREQUENCIES (IN CPS) MATCHED TO POSITIVE AND NEGATIVE SQUARE WAVES UNDER TWO LIGHT CONDITIONS

Square waves	.01 mL		Dark adapted	
pps	Positive	Negative	Positive	Negative
5	5.6	5.6	4.9	5.1
10	16.4	17.2	11.2	12.4
20	19.4	19.8	20.2	20.1

2. *Varying Frequency in Terms of cps*

In this experiment we matched positive, alternating and biphasic square waves of 4 ms pulse duration to sine waves over the frequency range 5 to 30 cps. As the comparison here is made in terms of cps, alternating and biphasic square waves have twice as many 4 ms pulses per second as do positive square waves. The result of this comparison is shown in Figure 1.

It is seen that in the dark adapted state the positive pulses are matched quite closely to the corresponding sine wave frequency over the whole frequency range. Thus the relationship is essentially a linear one. The biphasic pulses are matched to a higher frequency than the positive pulses at 5 and 10 cps, while for 15 cps and above, the biphasic pulses are matched to lower sine wave frequencies than are the positive pulses. In the low frequency range, below 15 cps, the alternating square waves are matched to higher sine wave frequencies than are the biphasic square waves. For higher frequencies there is little apparent difference between alternating and positive pulses, indicating that one of the polarities in the alternating square wave is not seen as a flash.

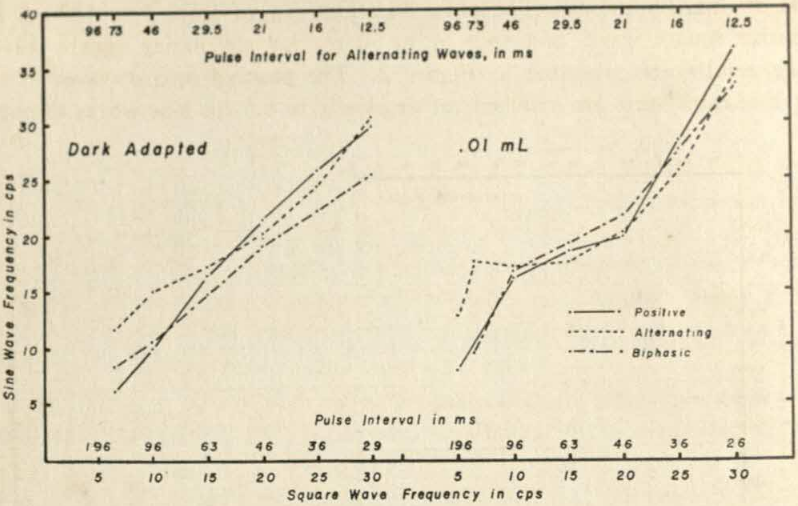


FIGURE 1
SINE WAVE FREQUENCIES MATCHED TO SQUARE WAVES
VARYING CPS (PULSE DURATION 4 MS)
Note: one cps alternating and biphasic square wave contains 2 pps.

In the .01 mL adaptation state all wave forms are matched to a higher frequency than in the dark adapted state, except at 20 cps. At the low frequencies the alternating pulses are matched to higher frequencies than are the positive or biphasic pulses. Above 10 cps the matches of all three wave forms are approximately the same.

The essential conclusion of this experiment is that at frequencies above 15 cps, alternating square waves, with two pulses per cycle, are matched to the same sine wave frequency as are positive square waves with only one pulse per cycle. At low frequencies alternating square waves are seen approximately twice as fast as monophasic square waves, apparently because both phases of the cycle are seen.[3]

3. *Varying Pulse Duration*

This experiment was concerned with the problem of whether or not a sudden change in matching occurs due to a make and break-effect as pulse duration increases. At a constant frequency of 6.5 cps, pulse duration was varied for positive and alternating square waves for dark adapted, .01 mL and 30 mL luminosity. The pulse durations ranged from 4 to 130 ms for positive square waves and from 4 to 70 ms for alternating square waves. The results are presented in Figure 2. The positive square waves in the dark adapted state are matched rather closely to 6.5 cps sine waves throughout the range of pulse durations. Positive square waves at .01 mL are matched to about 10 cps for pulse durations below 55 ms, from which point they are matched to increasingly higher frequencies reaching a maximum of 15 cps at 80 ms duration. At longer durations they are matched to increasingly slower frequencies, returning to a match of 10 cps at 125 ms duration. Alternating square waves, dark adapted, from 4 to 20 ms pulse duration, are matched to about 13 cps, twice the frequency to which the dark adapted positive pulses were matched. The match decreases gradually until at 60 ms duration it levels out at 6.5 cps. Finally, the alternating square waves at .01 mL start out with the highest match of them all, definitely higher than alternating square waves dark adapted, and higher than positive square waves at .01 mL. This high match is gradually reduced,

---

[3] That each phase of the alternating square wave may be seen as a flash has been corroborated by counting the number of flashes per second at frequencies of 1 and 2 cps.

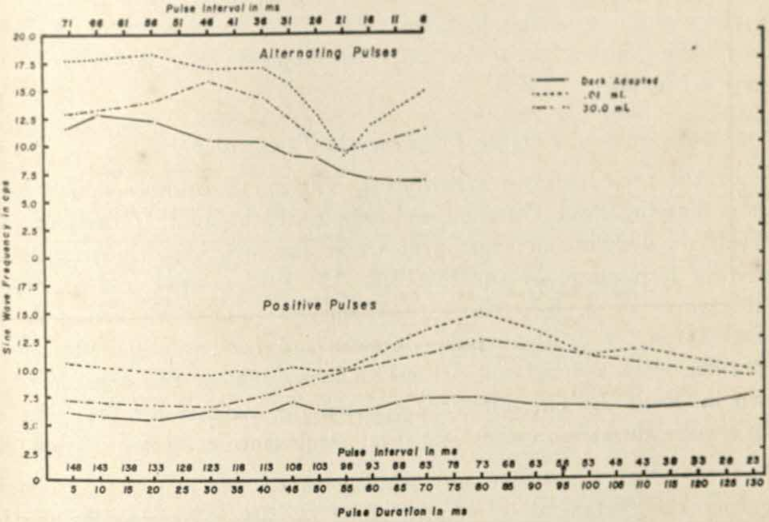


FIGURE 2
SINE WAVE FREQUENCIES MATCHED TO SQUARE WAVES
VARYING PULSE DURATION (FREQUENCY 6.5 CPS)
Note: one cps alternating square wave contains 2 pps.

reaches a minimum of 9 cps at 55 ms and rises rather sharply again. Thus, 55 ms duration seems to be a crucial point for both types of square waves when adapted to .01 mL.

Since a discrepancy appeared between the two light conditions, we decided to include a light condition of 30 mL. These data are presented in Figure 2 but will be dealt with below under the heading Varying Adaptation State.

Pulse duration does not seem to influence the flicker rate from positive square waves during dark adaptation. At .01 mL luminosity and for alternating square waves, the pulse duration influences the flicker rate in a differential manner.

4. *Varying Frequency in Terms of pps*

Since the preceding two experiments, Varying Frequency in Terms of cps and Varying Pulse Duration, are complicated by another variable, pulse interval, we designed an experiment where this was kept constant for the two wave forms, e.g., by equating pps. We used positive and alternating square waves of 5, 6.5, 10, 13, 15 and 20 pps. Each condition was replicated for 4 and 20 ms pulse duration and for two different states of adaptation: dark adapted and .01 mL. The results of this experiment are shown in Table 2. The difference between the matches for the 4 ms and 20 ms pulse duration reaches statistical significance only for 20 pps (.05 level of confidence). The 20 ms pulse duration contributes heavily to this significance. It appears, therefore, that this particular situation may be explained by insufficient recovery time between the pulses (see discussion below). It is worth noting, however, that there is a tendency for the 4 ms pulses to be matched to higher sine wave frequencies then the 20 ms pulses. Based on the lack of significance between the two pulse durations, we have combined the data from the 4 and 20 ms duration. The resulting

TABLE 2
SINE WAVE FREQUENCIES (IN CPS) MATCHED TO POSITIVE AND ALTERNATING SQUARE WAVES VARYING PPS UNDER TWO LIGHT CONDITIONS AND TWO PULSE DURATIONS

Square waves	Positive				Alternating			
	Dark adapted		.01 mL		Dark adapted		.01 mL	
pps	4 ms	20 ms	4 ms	20 ms	4 ms	20 ms	4 ms	20 ms
5	4.1	4.9	7.6	5.6	3.8	3.1	5.0	4.6
6.5	7.0	5.5	10.5	9.7	5.3	5.2	7.2	7.5
10	10.0	11.2	15.9	16.4	11.2	10.8	12.7	13.4
13	14.2	13.4	15.8	13.7	12.9	12.4	17.7	18.3
15	16.5	17.7	15.5	15.2	12.8	11.9	18.4	17.8
20	21.5	20.2	19.7	19.4	15.1	14.1	17.3	12.8

curves are presented in Figure 3. Again, it is seen that the matching of the positive pulses in the dark adapted state follows the "true" match very closely. The alternating pulses under the same conditions are matched to essentially the same frequency as are the positive pulses. Above 10 pps a discrepancy occurs which increases with increasing frequency.

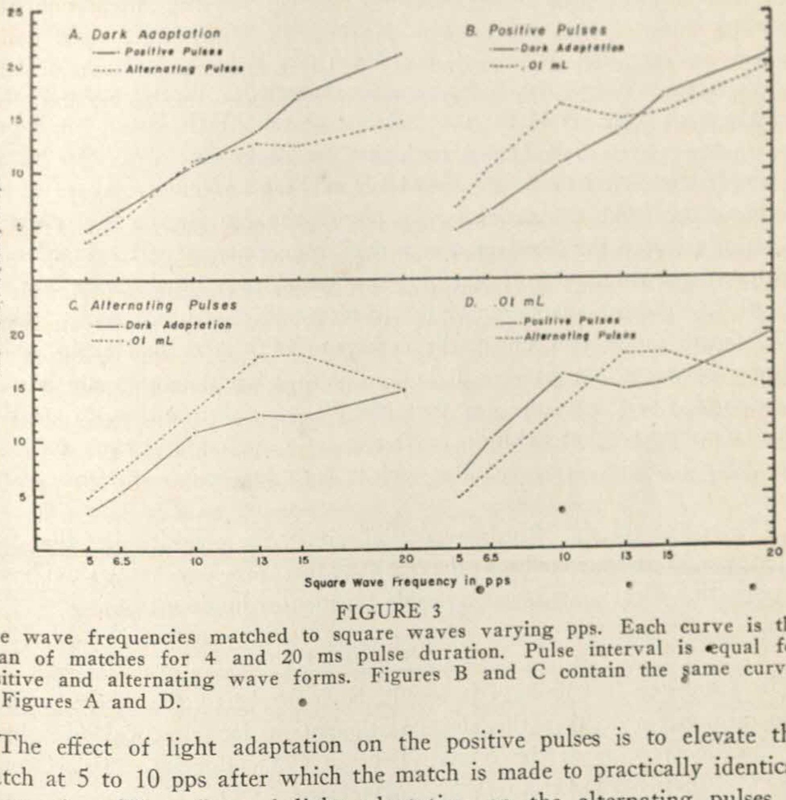


FIGURE 3
Sine wave frequencies matched to square waves varying pps. Each curve is the mean of matches for 4 and 20 ms pulse duration. Pulse interval is equal for positive and alternating wave forms. Figures B and C contain the same curves as Figures A and D.

The effect of light adaptation on the positive pulses is to elevate the match at 5 to 10 pps after which the match is made to practically identical frequencies. The effect of light adaptation on the alternating pulses is also to elevate the match. While the dark adapted curve levels off at 10 pps, the light adapted curve levels off from 13 to 15 pps and then decreases. This results in the biggest discrepancy at 13 and 15 pps rather than at 5 and 10 pps as seen in the positive curves.

Comparing the curves for positive and alternating pulses in the light adapted state, it is seen that they have some similar features, an initial increase followed by a plateau section, which occurs at slightly different frequencies. This is different from what happened in the dark adapted state, where a levelling off occurred for the alternating waves but not for the positive waves.

Thus it appears that repetition rate differentially affects flicker rate resulting from positive and alternating waves of equal interval between pulses, except for the dark adapted state below 10 pps.

5. *Varying Adaptation State*

We have consistently found that approximately veridical matching occurs for the positive waves in the dark adapted state. At the .01 mL condition the square waves are matched to higher frequency sine waves. This fact is amply demonstrated in Figures 1, 2 and 3. To understand better the influence of light adaptation we repeated the Varying Pulse Duration experiments with the *S*s adapted to a third luminosity, 30 mL. In Figure 2 are presented the results of matching sine waves to 6.5 cps square waves at the three different adaptation states (dark adapted, .01 mL and 30 mL). The figure shows that under the conditions of bright light the match is always between that of dim light and dark adaptation. Furthermore it appears that at short pulse durations the 30 mL curve is close to the dark adaptation curve, but with increasing pulse duration it draws closer to the .01 mL curve. This trend is present both for positive and alternating pulses.

The results for the 30 mL adaptation are in the opposite direction from what would be expected on the basis of the previous results. Thus it appears that flicker rate is not a simple function of luminosity.

C. DISCUSSION

In discussing our findings it should be remembered that there is fairly general agreement that the site of stimulation for electrical phosphenes is the bipolar cells of the retina (Bouman 1935, Bouman 1940, Clausen 1955, Meyer-Schwickerath 1951, Motokawa 1949, Motokawa, Iwama and Endo 1948, Rohracher 1935, Schaefer 1942, Schwartz 1940, Svaetichin 1956). The possibility exists, however, that several structures are simultaneously stimulated at current intensities well above threshold values. Whichever structures are stimulated, it stands to reason that the intensities utilized in our matching experiments will cause all stimulated structures to fire simultaneously rather than randomly. This is probably the most important feature which differentiates our results from results obtained with photic stimulation.

The first question which arises is what conclusion concerning flicker rate may be made from our technique of matching square wave stimuli to sine wave stimuli. Does this technique only allow us to compare relative flicker rate from the two types of stimuli, or does it allow us to make inference with regard to the actual flicker rate of the two stimuli? To answer this question we have utilized beats as described by Clausen and Vanderbilt (1957) as a means of determining the actual flicker rate resulting from sine waves. Starting with sine wave stimulation of such low frequency that flicker rate could be counted, we have produced beats by simultaneous stimulation with intermittent light of a frequency difference of 1 cps throughout the stimulus range of 2 to 20 cps. From this and other findings we have tentatively concluded that, at least up to 20 cps, the flicker rate is the same as the stimulus frequency of the sine wave current. If the data and the details of this argument, which are to be presented in a subsequent article, require another conclusion, this should not be of consequence for the basic principles listed below. It may be necessary, however, to conceive of them in terms of relative rather than absolute flicker rates.

Unlike the situation for square wave stimulation we have observed that the flicker rate from sine wave stimulation remains the same whether the eyes are dark or light adapted. Thus we think that, under all light conditions, the sine wave frequency to which a square wave is matched indicates directly the flicker rate ensuing from the square wave stimulation.

Monophasic square waves of short pulse duration, repetition rate up to 20 pps, in a dark adapted state, are matched to sine waves of corresponding frequencies, i.e., veridically matched. This direct relationship breaks down under certain circumstances with square waves of alternating polarities, long pulse durations, short intervals or exposure of the eye to light.

Although some of our findings remain unexplained, we think that the following principles are operating to produce the results we have recorded:

Principle 1: Following an electrical stimulus a *recovery period* is necessary before the nerve fibers are ready to respond again. This recovery period is shorter when two successive stimuli are of the same polarity than when they are of opposite polarity.

Principle 2: When the stimulus duration is above a certain value, a flash is seen at the beginning and at the cessation of the stimulus, i.e., a *make and break-effect*. Below a critical value the make and break-effect tend to merge. Similarly, when the interval between two pulses is sufficiently short, merging of a break-effect and the succesive make-effect occurs. Under certain circumstances, in the dark adapted eye, the make and break-effect is not perceived. In the dark adapted eye stimulation with a single DC pulse or with repetitive stimulation of extremely long pulse durations and intervals, make and break is seen. Within the range of pulse durations and corresponding pulse intervals used in the present study we have found that make and break are not perceived separately by the dark adapted eye. The difference in make and break-effect between the dark adapted and light adapted eye with repetitive stimulation remains a puzzle. As possible explanations we have considered recovery time, merging, grouping of light flashes, brightness contrast, polarization from photic stimulation, but neither of these explanations seem to fit the data.

Principle 3: Some of our data indicate a faster flicker rate than the stimulus frequency, which cannot be wholly accounted for by make and break. An additional mechanism of facilitation seems to be operating. This phenomenon may be a case of brightness contrast which occurs in a wide variety of perceptual situations. In the present situation a possible mechanism may be stated in terms of the *hyperexcitability* which is a part of the relative refractory phase of the nerve (Lloyd 1949). An electrical impulse throws the stimulated structures into complete refractoriness followed by a period of hyperexcitability which gives rise to an additional flash. The effect of this is of particular significance for the light adapted eye. When the eye is adapted to .01 mL, which is a rather dim light, the electrical stimulus gives rise to the sensation of a light flash because the luminous intensity of the phosphenes is higher than the photic background stimulation. The refractory period blocks momentarily the stream of impulses caused by the photic background stimulation. This in turn is followed by the period of hyperexcitability when the photic background stimulation is momentarily seen as more intense. This heightened intensity results from the fact that all nerve fibers are at this time in a hyperexcitable state and thus the system is capable of carrying more impulses simultaneously than is the case before the electrical stimulation when the fibers are firing randomly. This is analogous to the dark adapted eye being stimulated by a sudden flash of light.

In the dark adapted eye there is no background stimulation to take advantage of the heightened sensitivity.

Principle 4: When the eye is adapted to 30 mL, which is a rather bright light, the electrical impulse gives rise to a sensation of *interrupted light* because the luminous intensity of the phosphenes is lower than the photic background stimulation. The dark sensation which is predominant may be due to the absolute refractory period which blocks momentarily the photic background stimulation. Under these circumstances one cannot discriminate as a separate flash the period of hyperexcitability. When the pulse duration is sufficiently long, the dark sensation appears for break as well as for make-effect.

Principle 5: In varying such parameters as pulse interval or pulse duration we have found gradual rather than sudden changes in frequency matches. This may be explained on the basis of *probability*. Suppose we are concerned with the pulse duration which would respond at make and break. Below a certain duration a flash would occur at make but not at break. At a certain critical value, break responses would occur occasionally, and at increasing durations they would occur more and more regularly until a rate of twice the stimulus frequency would be seen. We assume this principle to work for recovery period and hyperexcitability as well as make and break-effect.

According to this explanation the time interval between the flashes at any given repetition rate may not be constant, but the overall rate is sufficiently high to obscure such irregularities.

Applying these principles to sine wave stimulation we see that the flicker rate is not affected by light adaptation, as noted above, because the waxing and the waning of the stimulus is too gradual for the refractory period to block all fibers simultaneously, and thus for the hyperexcitability to enhance the background light (Principle 3). The same reasoning excludes make and break-effect (Principle 2). Also there is not sufficient recovery time because of the nature of the stimulus, (Principle 1) to make the two phases of the cycle perceivable. Sine waves, therefore, act as a monophasic stimulus.

The findings presented in Figure 2 may now be discussed in terms of these principles. In the dark adapted state positive pulses are matched to about the veridical sine wave frequencies throughout the range of pulse durations. This is explained by the fact that in the dark adapted eye make and break (Principle 2) under the conditions of our experiment are absent. Also, recovery period (Principle 1) for positive pulses in the dark adapted eye is short enough to let each pulse be perceived, even at pulse intervals of about 25 ms.

Positive pulses at .01 mL have, at short pulse durations, no distinguishable break-effect. They are matched to about twice the rate of positive pulses in dark adaptation because of the action of hyperexcitability as explained in Principle 3. The curve is essentially flat to 55 ms pulse duration where break-effects begin to occur, with a resulting rise of the curve. This rise reaches its peak at the point where pulse duration and pulse interval are approximately equal, 80 ms and 73 ms respectively. We have found from other data that the hyperexcitability phase disappears in this region at intervals of 46-70 ms, in accordance with Principle 5. From this point (80 ms duration and 73 ms interval) the interval becomes too short for the break and the following make to be discernible. The probability of such merging of break and make becomes greater as the pulse interval becomes shorter, resulting in a progressive decrease of the curve.

The curve for the 30 mL positive pulses is matched to approximately 6.5 cps because of the mechanism described in Principle 4. At 30 ms pulse duration the curve begins to rise because there is now a small probability of seeing the break as well as the make. From this point make and break function in the same manner as described for the .01 mL curve. It should be noted that throughout the course of this curve hyperexcitability does not give rise to an additional flash, as stated in Principle 4.

With regard to alternating pulses presented in Figure 2 the situation is similar. Each positive and negative pulse is perceived because of the sudden onset of stimulation inherent in the square wave and the long interval between pulses. Thus alternating pulses in dark adaptation are seen about twice the frequency of the positive pulses in dark adaptation. When the pulse interval decreases below 56 ms, the recovery period, Principle 1, is too short to allow consistent perception of each pulse and the curve decreases. At about 20 ms one of the pulses becomes completely ineffective and the alternating pulses are now functionally perceived as monophasic ones, resulting in a match to 6.5 cps. Whereas we found that recovery time for positive pulses was less than 25 ms in the dark adapted state, we see here that recovery time in excess of 50 ms is necessary for subsequent stimuli of opposite polarity to be effective.

The flicker rate for alternating pulses at .01 mL is determined by Principles 1, 2 and 3. At short pulse durations and long intervals, recovery time is sufficient for every pulse to be seen. The rate is further enhanced by hyperexcitability. Pulse intervals, however, are approaching the critical value of 50-70 ms. The enhancement in flicker rate from 13 to 17 cps may be more than can be accounted for by hyperexcitability. We know of no other mechanism, however, which explains an increase in flicker rate over the stimulus frequency, which would work in conjunction with hyperexcitability to enhance the flicker rate.

When pulse duration is longer than 20 ms (pulse interval shorter than 56 ms) recovery time is too short for alternating pulses to be seen each time. Pulse interval is becoming increasingly less adequate to restore the nerve fibers to their original level of excitability in accordance with the probability principle. Around 30 ms pulse duration, the make and break-effect begins to operate. This effect, however, is interacting with recovery time which at this section of the curve is dominant, resulting in a reduction in flicker rate.

At pulse durations above 55 ms, interval 20 ms, recovery period has reached its critical value so that only one phase of the alternating pulses are seen. With the alternating pulses functioning as monophasic waves the break is allowed to reappear for the single pulse, since the make of the alternate pulse is no longer functionally present to merge with the break of the previous pulse. As pulse duration lengthens, increasing the probability of the break-effect, the curve takes a sharp rise. Thus the alternating curve in this region follows the positive .01 mL curve very closely.

Alternating pulses from 5 to 20 ms duration at 30 mL are seen at approximately the same rate as during dark adaptation since the flicker is the result of interruption of the background light which precludes hyperexcitability effect (Principle 4, see explanation for the corresponding section of the positive 30 mL curve). From this point the curve is determined by the same factors which are operating in the alternating .01 mL curve for corresponding pulse durations: namely recovery time, make and break, and probability (Principles 1, 3 and 5). It should be noted here again that after recovery time has reached its critical value, the alternating curve coincides with the positive curve for the same light condition: the alternating current is functionally acting as a monophasic wave.

The flicker rate of the alternating square waves in light adaptation is slightly higher than would be expected from our principles. Thus it is possible that some other mechanism may be operating to enhance flicker rate. The intrinsic periodicity of the visual structures as discussed by Granit (1955) and Bartley (1958, 1959) may be interacting with the stimulating frequency to create enhanced flicker rate, but we have not been able to comprehend how such an interaction would take place.

The interpretation above is, in essential parts, corroborated by the results presented in Figure 1. In dark adaptation the positive pulses are matched veridically throughout the frequency range; recovery time sufficient, make and break principle not operating and hyperexcitability not feasible. Alternating pulses in dark adaptation are seen twice as fast as positive pulses at 6.5 cps, i.e., with pulse interval of 73 ms full recovery is achieved and both phases of the cycle are seen (Principle 1). Somewhere between 73 and 46 ms recovery time becomes partially insufficient, and somewhere between 29 and 21 ms one phase is not seen at all. The corresponding values determined from Figure 2 were 56 and 20 ms. At pulse intervals below 20 ms recovery becomes inadequate and alternating waves act as monophasic stimuli.

The data for .01 mL in Figure 1 show that all three wave forms give rise to increased flicker rate as compared to the dark adapted state. We interpret this shift in the positive curve to be a result of hyperexcitability. As frequency increases, pulse interval decreases, resulting in a decrease in probability of the hyperexcitability phase being seen as an additional flash. Somewhere between pulse intervals of 96 and 62 ms (10 to 15 cps) this decrease begins to take place and between 62 and 46 ms (15 to 20 cps) the hyperexcitability ceases to be a factor and the match becomes veridical. At intervals shorter than 46 ms (frequencies above 20 cps) the flicker rate is again raised by some factor we have been unable to comprehend.

At low frequencies (5 and 6.5 cps) the curve for alternating pulses is determined by sufficient recovery time and hyperexcitability (Principles 1 and 3). Hyperexcitability disappears somewhere below 10 cps (72-46 ms interval). The curve, running parallel to the positive curve, is still raised, because the recovery time is still operating sufficiently to see the alternate phase. At 20 cps (interval of 21 ms) recovery time has become inadequate, so that only one phase of the cycle is seen. At 30 cps there is an increase in perceived rate as there was in the positive .01 mL curve, which remains unexplained.

As will have been noticed there is close agreement between data from Figures 1 and 2 in the values determined for recovery time and hyperexcitability phase.

It seems consistent with our explanation that results from biphasic stimulation at both adaptation levels are essentially like those from positive stimulation since there is no interval between one of the phase reversals; there is no recovery period and the second phase (the negative) remains inactive.

Turning now to the data presented in Figure 3A and Table 2, we may explain the difference between positive pulses, which are veridically matched, and alternating pulses in dark adaptation by Principle 1. Flicker rate from alternating pulses begins to fall below that from positive pulses at 13 pps because at this point the recovery period is beginning to be insufficient for alternating pulses. The pulse interval where this takes place is between 72 and 56 ms.

Positive pulses at .01 mL in the range below 13 pps produce a faster flicker rate than those at dark adaptation as a result of the hyperexcitability effect. From the data in Figure 3B the interval necessary for its functioning may again be determined. An interval longer than 72-80 ms seems to be necessary for maximal operation of this factor. At intervals shorter than 52-72 ms hyperexcitability can no longer function and the match is a veridical one as it is during dark adaptation.

In comparing alternating pulses for the two light conditions (Figure 3C) it appears that the light adapted condition (.01 mL) gives rise to a higher flicker rate (Principle 3). Hyperexcitability begins to fail at about the same time as recovery time begins to become inadequate for stimulation by the alternating phase. At this point (13 pps), the perceived rate begins to decrease. This transition is completed at 20 pps where the two curves coincide.

The light adapted alternating curve is somewhat troublesome to our conceptual framework. This is specifically indicated in Figure 3D where it is difficult for us to explain why the alternating pulses are matched to lower sine wave frequencies than are the positive pulses at 5-10 pps.

We have considered other possible explanations for these findings, such as differential effects of on and off fibers, but have found this untenable since all neural structures in our stimulus situation must be activated simultaneously. Neither do we think that the resting potentials on the eye itself which have been discussed by Granit (1955) can contribute to an explanation, in that monophasic pulses of positive and negative polarity make no difference with regard to perceived flicker rate.

In reviewing the various determinations our data have allowed for the time courses of the underlying mechanisms for Principles 1, 2 and 3, we find the following over-all values. Recovery time for alternating pulses has a range of 20-60 ms, where the lower value indicates a mean interval too short for recovery and the upper value a mean interval beyond which recovery occurs consistently. Concerning positive pulses our data only allow us to say that an interval as short as 25 ms is sufficient for complete recovery in a dark adapted eye, thus definitely indicating that recovery period for alternating pulses is longer than that for positive pulses. The mean values for make and break are 30 and 80 ms, where the lower value indicates the minimum pulse duration for break effect and the upper value indicates pulse duration where the break effect is seen consistently.[4] The mean values of the critical range for hyperexcitability are 60-80 ms. Below 60 ms hyperexcitability does not occur at all, and above 80 ms it occurs regularly. Although our values are not precise, the close agreement between the various determinations gives a reasonable amount of validity to them.

The values for hyperexcitability are well within the range of 50-80 ms which Lloyd (1949) quotes from Grundfest as the negative after-potentials in C-fibers. We think it is justified to regard the bipolar cells in the category of C-fibers.

Our findings have established beyond doubt that sine waves and square waves have totally different stimulating properties. As sine waves have a more constant mode of action under various stimulus conditions, they may, in many instances, be the preferred stimulus source. Square waves, being less homogeneous in their action necessitate a more complicated design in order to produce unambiguous results.

We think that the five principles we have used to explain our findings are the major factors operating to produce the recorded results. When not all of the results are explained by these principles we think that it is a matter of finding additional factors rather than finding substitutes for those listed. Most of the difficulties occur in situations where hyperexcitability is used as an explanatory principle. It may be that there are factors other than light adaptation and pulse interval that influence hyperexcitability; closer scrutiny of the effect of polarity of the stimulating current and intrinsic periodicity of the system may provide a lead. Further investigation of the parameters influencing the flicker rate at 55 ms pulse duration (see Figure 2), may also be fruitful. In the meantime it should be a routine matter to design experiments to verify the factors which we have postulated and to determine more precisely the values which the present study has indicated.

D. Summary

A series of experiments have been performed in which the rate of electrical phosphenes produced by square wave stimulation has been matched to

---

[4] The 30 ms value corresponds to Howarth's (1954) minimal value of 30 ms in the strength duration curve, which he explains in terms of maximal summation of the on and off responses. Since it is questionable that on and off mechanisms are involved in electrical stimulation, we prefer to use the terms make and break in this connection, as there is little evidence that these processes are the same as those aroused by light. We assume that in electrical stimulation the current is acting upon nerve fibers whose electrical properties are identical whether they are connected to on or off elements.

phosphenes produced by sine wave stimulation. Such a comparison has been made for positive square and alternating square waves where repetition rate, pulse duration and light condition have been varied.

Introductory experiments indicated that there is no difference between positive and negative square waves with regard to flicker rate. The major findings are as follows:

1. At low frequencies of alternating square waves, both the positive and the negative phase of the cycle is seen as a flash. At frequencies above 15 cps only one phase of the cycle is seen.

2. Varying pulse duration does not modify the flicker rate when the dark adapted eye is stimulated by positive square waves. When the eye is adapted to light, pulse duration does modify the flicker rate resulting from both positive and alternating square waves.

3. In general it is found that varying repetition rate affects the flicker rate from positive pulses. State of light adaptation modifies these repetition rate effects.

4. Comparison of three adaptation states shows consistently higher matches for a low light level. A high light level produces matches which fall between those for dark adaptation and the low light level.

A visual beat frequency technique is cited in support of the position that the sine wave frequency to which a square wave is matched, indicates directly the flicker rate of the square wave.

To explain the recorded results we have employed the following principles:

1. A recovery period (which is different for positive and alternating square waves) is necessary before the nerve fibers are ready to respond to a subsequent stimulus.

2. When pulse duration reaches a certain value a break-effect occurs in addition to the make-effect, resulting in enhanced flicker rate. In the dark adapted eye the break-effect does not occur with repetitive stimulation, except at extremely low frequencies.

3. An electrical stimulus results in total refraction, followed by a period of hyperexcitability during which the photic background stimulation is momentarily seen as more intense. This is an additional mechanism for enhancement of flicker rate.

4. When the eye is adapted to a relatively bright light an electrical stimulus gives rise to the sensation of interrupted light rather than light

flashes. This may be due to the absolute refractory period momentarily blocking the photic background stimulation.

5. Principles 1, 2 and 3 operate on a probability basis. Under certain conditions they occur at each pulse, whereas under different conditions they occur only part of the time, which explains the gradual changes which occur when one of the above parameters is systematically varied.

Some of our findings cannot be explained by the above principles. We think that additional principles are needed rather than substitutes for those listed.

REFERENCES

1. BARGER, D. M., & GEBHARD, J. W. An electronic stimulator for use on the human eye. *Amer. J. Psychol.*, 1952, **65**, 450-457.
2. BARTLEY, S. H. Central mechanisms of vision. *In* J. Field, H. W. Magoun and V. E. Hall (*Eds.*), *Handbook of Physiology*, Section I: *Neurophysiology*. Vol. I. Washington, D. C.: *Amer. Physiol. Soc.*, 1959. Pp. 713-740.
3. ———. Some factors and concepts regarding the neurophysiology of the optic pathway. *AMA Arch. Ophthal.*, 1958, **60**, Part II, 775-791.
4. BOUMAN, H. D. Experiments on the electrical excitability of the eye. *Arch. Néerl. Physiol.*, 1935, **20**, 430-445.
5. ———. Electrical excitability of the eye: A study in clinical physiology. *Ophthalmologica*, 1940, **99**, 394-401.
6. CLAUSEN, J. Visual Sensations (Phosphenes) Produced by AC Sine Wave Stimulation. Copenhagen: Ejnar Munksgaard, 1955.
7. CLAUSEN, J., & VANDERBILT, C. Visual beats caused by simultaneous electrical and photic stimulation. *Amer. J. Psychol.*, 1957, **70**, 577-585.
8. GEBHARD, J. W., DUFFY, M. M., MOWBRAY, G. H., & BYHAM, C. L. Visual sensitivity to the rate of electrically produced intermittance. *J. Opt. Soc. Amer.*, 1956, **46**, 851-860.
9. GRANIT, R. Receptors and Sensory Perception. New Haven: Yale University Press, 1955.
10. HOWARTH, C. I. Strength duration curves for electrical stimulation of the human eye. *Quart. J. Exp. Psychol.*, 1954, **7**, 47-61.
11. LLOYD, D. P. C. Functional property of neurons. *In* J. E. Fulton (*Ed.*), *Textbook of Physiology*. (16th ed.) Philadelphia: W. B. Saunders, 1949. Pp. 33-57.
12. MEYER-SCHWICKERATH, G., & MAGUN, R. Ueber selektive elektrische Erregbarkeit verschiedener Netzhautanteile. *Arch. Ophthal.*, 1951, **151**, 693-700.
13. MOTOKAWA, K. Visual function and the electrical excitability of the retina. *Tohoku J. Exp. Med.*, 1949, **51**, 145-153.
14. MOTOKAWA, K., IWAMA, K., & ENDO, T. Ueber den Einfluss unterschwelliger elektrischer Reizung des Auges auf die Lichtempfindlichkeit der Netzhaut. *Tohoku J. Exp. Med.*, 1948, **49**, 331-338.
15. ROHRACHER, H. Ueber subjektive Lichterscheinungen bei Reizung mit Wechselstroemen. *Z. Sinnesphysiol.*, 1935, **66**, 164-181.
16. SCHAEFER, H. *Elektrophysiologie*. Vol. II. Wien: Deuticke, 1942.
17. SCHWARZ, F. Quantitative Untersuchungen ueber die optische Wirkung sinusfoermiger Wechselstroeme. *Z. Sinnesphysiol.*, 1940, **69**, 1-26.
18. SVAETICHIN, G. Aspects of human photoreceptor mechanism. *Acta Physiol. Scand.*, 1956, **39**, Supplement **134**, 93-111.

19. WOODWORTH, R. S. Experimental Psychology. New York: Henry Holt, 1938.
20. YAEGASHI, S. The site of electrical excitation of the human eyes by the sinusoidal alternating currents at frequencies of 20 cps. *Tohoku J. Exp. Med.*, 1958, **67**, 163-172.

The Training School
Vineland, New Jersey

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 47-49.

THE RELATION BETWEEN THE ASSOCIATION VALUE, PRONUNCIABILITY, AND AFFECTIVITY OF NONSENSE SYLLABLES*

E. I. du Pont de Nemours & Co., Wilmington, Delaware

WARNER R. WILSON AND JAMES C. BECKNELL

Students of verbal learning have traditionally given much attention to the variables of association value and frequency and more recently, pronunciability. The dependent variable which has been studied almost exclusively is ease or speed of learning. The purpose of this study is to relate association value and pronunciability to affectivity. Pronunciability and association value have been repeatedly demonstrated to have a powerful influence on ease of learning although recent work suggests that pronunciability is the basic variable. This study was conducted to obtain information relative to the validity of a general hypothesis to the effect that the variables that influence learning also influence preference or choice behavior. Rankings, indicating how much the *S* liked certain nonsense syllables and how inclined she would be to buy a product with the syllables as a brand name, were interpreted as measures of affective value.

A. METHOD

1. *Subjects*

The *S*s were 39 secretaries employed by E. I. du Pont de Nemours & Company.[1]

2. *Materials*

Nine nonsense syllables of 100 per cent association value and nine nonsense syllables of 0 per cent association value were used (Glaze, 1928). An attempt was made to choose syllables that were not markedly similar to each other. Three lists of six syllables were constructed. Each of the three lists contained three syllables of high and three of low association value. Within each list high and low association value syllables followed each other in a randomly determined manner.

---

* Received in the Editorial Office on January 24, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The authors are indebted to their secretaries, Mary Biondi and Dorothy Callahan, who contacted the *S*s, typed the data sheets, and typed and corrected the manuscript.

3. *Procedure*

*S*s were instructed to rank each list of syllables under a different set of instructions. The instructions were to rank in terms of how easily the *S* could pronounce the syllables, how much she liked them, and how inclined she would be to purchase a product with each syllable as a brand name. Each *S* ranked the lists in the same order, but the order in which the different instructions were given was systematically randomized. Happily, there was no evidence that order had any effect on the ratings. In effect, each *S* rated three equivalent groups of syllables under three different instructions.

B. DISCUSSION AND RESULTS

The data are exceptionally clear-cut as was expected. One method of analysis employed involved adding the ranks received by the high and low association value syllables. When this analysis is done, practically every *S* supports the hypothesis of a relationship between association value and pronunciability and affectivity. In all but two cases, the high association value syllables have a higher average rank. The probability that this strong a relationship would occur by chance is of course far less than .001. The comparison of average rank, however, does not constitute a very stringent test and it seemed desirable to also determine the percentage of times the affectivity and pronunciability ratings were in complete agreement with the previously determined association values, that is, how often the high association syllables received the ranks of one, two, and three. It is obvious from this more stringent analysis that the association values do not predict the other ratings with perfect accuracy; it agrees with the first analysis, however, in that it demonstrates a strong relation between association value and pronunciability and affectivity. Also, these data are quite compatible with the hypothesis that the relationship with association value is the same or approximately the same whether the ratings are in terms of pronunciability, liking, or inclination to buy. Ratings on pronunciability agree perfectly with the association values 32 times out of 39 or 82 per cent of the time. The comparable figures for both liking and inclination to buy are 29 times out of 39, or 74 per cent of the time. The difference between 74 per cent and 82 per cent does not approach significance. These results lend firm support to the hypothesis that choice behavior is affected by the same variables that affect learning.

C. SUMMARY

It was predicted that a strong relationship could be demonstrated between the association value of nonsense syllables and ratings of their pronunciability and affectivity. Conclusive positive results were obtained. The theoretical implication is that both learning and choice behavior are affected by the same variables.

REFERENCE

1. GLAZE, J. A. The association value of nonsense syllables. *J. Genet. Psychol.*, 1928, **35**, 255-267.

Department of Psychology
University of Hawaii
Honolulu 14, Hawaii

Published as a separate and in *The Journal of Psychology*, 1961, 52, 51-53.

PHENOMENAL ORIENTATION AND PERCEIVED ACHROMATIC COLOR*[1]

Department of Psychology, University of Kansas

WILLIAM EPSTEIN

The explanatory problem relating to brightness-constancy stems from the absence of a univocal correspondence between local retinal intensity and perceived color. The intensity of the retinal region corresponding to any given distal stimulus will vary with changes in the level of illumination. Despite these wide variations in retinal intensity perceived achromatic color will often remain unchanged. On other occasions, conditions of illumination may bring about a situation in which targets of different albedo produce identical retinal intensities. Under these conditions the targets will usually be correctly perceived, *i.e.*, perceived color will approximate albedo requirements.

The explanation of these phenomena which is generally accepted attributes constancy to an unconscious correction for conditions of illumination (4, pp. 428-444). However, Wallach (3) has proposed an alternative interpretation. Wallach maintains that the proximal ambiguity mentioned above is not present if the stimulus-situation is defined relationally. He reports evidence purporting to show that perceived achromatic color is determined entirely by an interaction of the intensities of the neighboring retinal regions corresponding to a target and the immediately adjacent surrounding. Since intensity-ratios are invariant for changes of illumination which affect the entire field, constancy is assured.

Recently, Hochberg and Beck (2) have argued that Wallach's intensity-ratio theory is insufficient. In an ingenious experiment Hochberg and Beck show that perceived achromatic color will vary with changes in the *phenomenal* orientation of the target although the intensity-ratio remains unchanged by variations in apparent slant.[2] In the present note a similar experiment will be reported.

---

* Received in the Editorial Office on January 25, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The effects described in this note were observed initially as incidental effects in a study of the slant-shape invariance hypothesis. The work was supported by a grant (M-4153) from the U.S. Public Health Service.

[2] Hochberg and Beck object to Wallach's thesis for the same reason that the Gestalt psychologists objected to the "constancy thesis." Indeed, Wallach's view might be considered to be a restatement of the constancy thesis in relational terms.

The apparatus consisted of a completely enclosed rectangular tunnel 7 feet in length. An aperture for monocular vision was present in the front wall along with a device for fixing *O*'s head at the desired reference-point. On the inside of the front wall was a circular fluorescent light. Six feet from the observation-aperture, near the end of the tunnel, and black and white checkerboard panel was inserted. The angle of orientation of this panel could be varied at will. In the present demonstration the panel was tilted 20° from the perpendicular away from *O*. An isoceles triangle (10 in.—base, 8 in.—altitude) cut out of textureless white cardboard was mounted in the objective center of the panel at a 45°—angle of slant outward from the apex of the triangle. The demonstration depends on the observation "that a flat shape of homogenous color will, when slanted away from a visible textured background, assume the slant of its background if binocular disparity and motion perspective are eliminated from stimulation" (1, pp. 130-131). In the present case the background induces an error in the perception of slant of the target. The target assumes the orientation of the background, *i.e.,* it appears to be tilted much less than its actual slant.[3]

Twenty *O*s viewed the target monocularly and with the head motionless. The *O*s were asked to report the perceived color of the target by selecting a matching color from a graded series of Munsell achromatic papers presented outside the tunnel in normal illumination. Ten *O*s chose Munsell No. 7 (albedo = 43.1 per cent), six *O*s selected Munsell No. 8 (albedo = 59.1 per cent) and four *O*s selected Munsell No. 6 (albedo = 30.0 per cent). The median choice was 7.0 (Q = 0.75).

Immediately following the monocular match *O*'s head was released to allow motion perspective, and the aperture was modified to permit binocular vision and restore retinal disparity. Under these conditions *O*s see a triangular target tilted away from the background, *i.e.,* perceived tilt now shifts in the direction of objective tilt.[4] The same *O*s were asked to match the color of the target again. The matches differed very little from the matches made under the reduced situation when slant was misperceived. Sixteen *O*s selected Munsell No. 7, two *O*s chose Munsell No. 6, and two *O*s selected Munsell No. 9 (albedo = 78.7 per cent). The median and Q were identical with those obtained before. Of the 20 *O*s, ten chose a slightly darker match when viewing the target binocularly, six chose a slightly lighter gray, and four *O*s made the same match under both conditions.

---

[3] The textural gradient is the stimulus for the perceived tilt of the background. It should be added that the target does not appear to *O* in its objective form but instead its appearance corresponds to the projective shape.

[4] For most *O*s there occurs also a change in perceived shape away from projective shape toward objective shape.

These results are precisely what is to be predicted from Wallach's thesis. Since the intensity-ratio remains unchanged by alterations in the perceived slant of the target, perceived color also should be unaffected. However, the results are not in line with the expectations suggested by Hochberg and Beck's study. Nor do the present results fit well with the traditional theory. Both of these latter points of view require that under conditions of unimpeded binocular viewing perceived color differ radically from the color perceived when the slant of the target is misperceived. In the latter case the target should appear gray as indeed it did. However, in the free-viewing situation with slant accurately perceived, color constancy should have been restored, *i.e., O* should have reported a white target or at least a color match which was considerably lighter than the initial monocular match. This change in perceived color did not occur.

The foregoing is not intended as a challenge to the reliability of Hochberg and Beck's findings. Essentially the same results were obtained by the present writer in an unreported replication of Hochberg and Beck's experiment. However, the observations reported in this note suggest that the Hochberg-Beck findings may be peculiar to the specific experimental operations required by their method of investigation.

REFERENCES

1. BECK, J., & GIBSON, J. J. The relation of apparent shape to apparent slant in the perception of objects. *J. Exp. Psychol.,* 1955, **50**, 125-133.
2. HOCHBERG, J. E., & BECK, J. Apparent spatial arrangement and perceived brightness. *J. Exp. Psychol.,* 1954, **47**, 263-266.
3. WALLACH, H. Brightness constancy and the nature of achromatic colors. *J. Exp. Psychol.,* 1948, **38**, 310-324.
4. WOODWORTH, R. S., & SCHLOSBERG, H. Experimental Psychology, rev. ed. New York: Holt, 1954.

Department of Psychology
University of Kansas
Lawrence, Kansas

Published as a separate and in *The Journal of Psychology*, 1961, 52, 55-61

VALUE ORIENTATIONS AND PREFERENCE FOR A MINIMAX STRATEGY*[1]

Department of Psychology, Ohio State University

ALVIN SCODEL

A. INTRODUCTION

Of late, experimental situations utilizing games have come to interest an increasing number of psychologists, but with few exceptions, such as the recent studies of Deutsch (1960) and Chaney and Vinacke (1960), the personality correlates of behavior in games have remained unstudied.

The present study is an attempt to elicit differences in value orientations between those who play a minimax strategy in a two-person, zero-sum game and those who play a riskier strategy.[2] Since a minimax strategy insures the least bad possible outcome against all possible choices of one's opponent, a strong preference for such behavior in a zero-sum situation can be equated with a disposition to adopt a defensive or conservative stance vis-à-vis one's opponent.

The decision to correlate value orientations with a preference for a minimax strategy was not gratuitous. In a study comparing the characteristics of high-accident automobile drivers with those of moderate- and no-accident groups, Conger, Gaskill, Glad, Rainey, Sawrey and Turrell (1957) found that the only discriminating measure in a large battery of tests was the Allport-Vernon-Lindzey Study of Values (somewhat modified in content to allow for less information and verbal facility in respondents). As compared to moderate- and no-accident drivers, high accident drivers were lower on the Religious and higher on the Theoretical and Aesthetic values. Scodel, Ratoosh and Minas (1959) found that subjects with a preference for high payoff-low probability outcomes in a situation involving decision making

---

* Received in the Editorial Office on January 30, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was supported in whole by the United States Air Force under Contract No. AF49 (638) -317 monitored by the Air Force Office of Scientific Research of the Air Research and Development Command.

[2] Zero-sum games are directly competitive games in which the sum of utility payments to the players is zero. In contrast, non-zero-sum games are games which may be played competitively but also afford players the possibility of collusion for mutual gain. Detailed accounts of zero-sum and non-zero-sum games may be found in Von Neumann and Morgenstern (1947) and Luce and Raiffa (1957).

under conditions of risk were significantly higher on the Theoretical value of the Allport-Vernon-Lindzey than subjects with a preference for low payoff-high probability outcomes. The difference in the Theoretical value, then, is consistent for two exceedingly different risk-taking situations (assuming, of course, that a large number of accidents is indicative of high risk taking). Risk taking in a zero-sum game represents still another kind of situation and if similar differences between high and low risk takers should emerge, the results would suggest a certain generality in risk-taking behavior by people of particular value orientations.

B. METHOD

Obviously, one can construct any number of zero-sum games. Several games were presented, and the one which was finally selected was chosen on the basis of its interest for players. The payoff matrix for the game is presented in Table 1.

TABLE 1
Payoff Matrix for the Game*

			Red	
		1	2	3
Blue	1	—2	15	1
	2	8	—6	—3
	3	3	3	2

* Positive entries represent payoffs in cents to the blue or row player; negative entries represent payoffs in cents to the red or column player.

The instructions to each pair of players were as follows:

> "I am going to give each of you $2.00 to play with in this game. You will both be allowed to keep the money you have at the end of the game. Here are three cards for each of you; an ace, a deuce and a tray. Now this is what you will do. See this board (*E* pointed to the matrix which was tacked on the wall in front of the *S*s). You (pointing to *S* on *E*'s left) will be the blue player and you (pointing to *S* on *E*'s right) will be the red player. On each trial blue will select one of the three rows and red will select one of the three columns. After you have selected your row or column, you will designate your choice by placing the appropriate card face down on the table. When both of you have placed the cards on the table, they will be turned up and the payoff will be determined by the intersection of the choices on this board. For example, if blue selects a deuce and red selects a deuce, blue gives red six cents; if blue selects an ace and red selects a tray, red gives blue one cent. (There were no negative numbers on the board. Payoffs to the blue player were in blue crayon and payoffs to the red player were in red crayon.)

> Now, I want you to look at the board for about two minutes. Then you will play the game 25 times. After that, you will reverse positions; blue will become the red player and red will become the blue player, and you will play the game 25 more times."

After the first half of the game, *S*s were given two more minutes to study the matrix before resuming the game in their reversed positions. The second half of the game was reversed for two reasons. First, not only do the entries in the matrix favor the row player, but the game's "saddle point" at 3-3 (3 is the minimax play for both players) has a value of two cents to the row player. By reversing roles the initial column player was given a chance to make up any incurred deficit. Second, it was anticipated that the initial blue player would almost invariably have more than the initial red player after the first 25 trials, and so choices in the second half afford an opportunity to study differences in game behavior as a function of being ahead of or behind one's opponent, quite apart from any relationships between values and a minimax strategy.

Thirty-five pairs of *S*s were run in the game. All *S*s were males enrolled in the introductory psychology course at Ohio State University. Each had been administered the Allport-Vernon-Lindzey Study of Values (3rd Ed.) prior to the game.

C. RESULTS

One possible result in such a game is that pairs will quickly adopt a minimax strategy and not deviate from it. If one player does not deviate from the 3 play, the other player has no choice but to play 3 also unless he has reason to believe that the first player will depart from 3. Actually, just one pair of the 35 fixated on a 3-3 strategy (which was maintained after the reversal of positions), and this pair was excluded from further consideration. The data to be presented refer to the remaining 34 pairs.

An incidental finding is that initial row players, all of whom had more money than their opponents at the end of the first 25 trials, play in a significantly more conservative manner when they become column players. The mean number of 3 plays increased from 10.50 to 14.71 (S.D.'s $= 4.36$ and 5.05, $t = 3.59$, $p < .01$). On the other hand, the change in 3 plays for initial column players from a mean of 9.57 to a mean of 11.40 was not significant (S.D.'s $= 3.98$ and 5.24, $t = 1.60$, $p > .10$). There was no significant difference in 3 plays for the first half of the game between the row and column players. A median test applied to the changes in 3 plays for *S*s of both groups produced a chi-square value of 3.80 ($p = .05$).

The initial row players converge more readily, then, on a minimax strategy.

Since initial row and initial column players differed significantly in their total number of 3 plays for the entire game, high and low 3 players were selected separately from each group and then pooled. The data on values, presented in Table 2, refer to 18 high and 18 low 3 players, with the nine highest and the nine lowest in the total 50-trial game taken from each of the two groups. The highs and lows represent, roughly, the upper and lower quartiles in number of minimax plays.

As indicated in Table 2, the low minimax players are higher on the Theoretical and lower on the Religious value.

TABLE 2
ALLPORT-VERNON-LINDZEY SCORES FOR THE HIGH AND LOW MINIMAX (3) PLAYERS

	Theoretical	Economic	Aesthetic	Social	Political	Religious
High (N = 18)						
Mean	41.94	40.83	33.05	34.95	44.78	45.00
SD	4.96	5.29	7.09	7.23	6.32	7.46
Low (N = 18)						
Mean	48.05	42.11	33.33	35.45	42.67	37.83
SD	4.72	7.06	5.91	5.81	6.58	9.44
t	3.62**	.53	.11	.22	.95	2.41*

* significant at .05 level
** significant at .01 level

D. DISCUSSION

Before discussing the value scores, some explanation should be given for the difference in behavior between the row and column players. Evidence exists (Lieberman, 1960) that in two-person, zero-sum games the number of minimax plays increases for both players with the number of trials, although roughly half of Lieberman's *S*s did not conform to the minimax model after 200 trials. It is quite possible that, in the present instance, initial column players would have started to play 3 more often if the game had persisted, but it is still of some interest that they did not converge on 3 as rapidly as the initial row players. Our tentative explanation is that, in a sense, the second half of the game is dominated by the initial row players who, now that they are ahead, can play cautiously and, by so doing, more or less insure their chances of having more money than their opponents at the end of the game.[3] Initial column players, on the other hand, might be inclined to play less cautiously in the second half of the game in order to obtain an edge over their opponents, but the increased 3 plays of their opponents could well force them into more 3 plays than they would otherwise make. Therefore, the absence of difference in 3 plays for the initial column players between the fiirst and second halves of the game can be explained as the result of a compromise between the wish to play less cautiously and an externally imposed, reluctantly undertaken defensive strategy.

---

[3] To insure more money than their opponents, repeated 3 plays in the second half of the game by initial row players would necessitate a margin over their opponents after the first 25 trials in excess of .50. Actually, 24 pairs had

Turning to the values, it is apparent that the differences between minimax players and non-minimax players are strikingly similar to those cited for low and high risk takers in the other studies. In all three studies low risk takers are lower on the Theoretical value, in two of the three they are higher on the Religious value, and in the study of low and high accident drivers they are lower on the Aesthetic value. (This difference in the Aesthetic value is in the same direction for both our risk-taking studies but not significant.)

Although it is patently unwise to draw large-scale generalizations from a single instrument like the Allport-Vernon-Lindzey, some interesting analogies can be drawn with the Kelly and Goldberg (1959) study of clinical psychologists. On the basis of a series of factor analyses they divided clinical psychologists into three groups of specialists: therapists, academicians, and administrators. Although therapists proved to be a highly heterogeneous group, the others turned out to be quite homogeneous. Among other things said of the academician he is . . ."theoretically oriented . . . . He has rejected religion—at least in its conventional forms . . . . His interests approximate more nearly those of scientists—perhaps in all fields—and artistic or creative persons in many areas" (p. 22). By way of contrast it is said of the administrator that "He was more likely to have some trouble with mathematics and science . . . . In college his interests and values were more conventional ones . . . . He was religious, in the sense that he attended church more regularly and considered himself a Protestant or Catholic (p. 23)."

The parallels between these two groups and the value differences obtaining between high and low risk takers are immediately apparent. What is of some interest is that for zero-sum games, game theory provides a normative model which would dictate the choice of a minimax strategy as the optimal

---

such a margin, but the explanation given here does not assume a completely worked-out strategy by the players. In support of our explanation, the correlation between amount ahead by initial row players after 25 trials and their number of 3 plays in the second half is .35, $p < .05$.

or "rational" strategy. It is possible, of course, to set up any number of criteria for defining "rationality" (cf. Garfinkel, 1960), but, in any case, theoretically oriented people—those whom one would expect to have scientific interests—tend to avoid the optimal or "rational" strategy. It may simply be, as Lieberman suggests on the basis of his interviews with *S*s who did not conform to the minimax model, that low minimax players are trying to avoid boredom. Certainly, the use of other choices does inject more interest and variety into the game, and from this point of view, it is not surprising that people with less conventional interests should play the game in this way.

A final though obvious point is that, inasmuch as some communality in value orientations has been demonstrated for high and low risk takers in very different kinds of risk-taking situations, the relevance of personality variables for theories of risk taking has been given additional plausibility.

E. Summary and Conclusions

Thirty-five pairs of subjects played a two-person, zero-sum-game with a saddle point. The game ran for 50 trials, and the positions were reversed after the first 25 trials. Each subject had been administered an Allport-Vernon-Lindzey prior to the game.

With respect to behavior in the game it was found that the group that was ahead of its opponents at the end of the first half of the game tended to converge more readily on a minimax strategy when the positions of the players were reversed.

Players who inclined toward a minimax strategy were lower on the Theoretical value and higher on the Religious value of the Allport-Vernon-Lindzey than players who avoided such a strategy. The relationships between this study and some other studies involving correlations between values and behavior were briefly discussed.

References

1. Chaney, M. V., & Vinacke, E. Achievement and nurturance in triads varying in power distribution. *J. Abnorm. Soc. Psychol.,* 1960, **60**, 175-181.
2. Conger, J. J., Gaskill, H. S., Glad, D. D., Rainey, R. V., Sawrey, W. L., & Turrell, E. S. Personal and interpersonal factors in motor vehicle accidents. *Amer. J. Psychiat.,* 1957, **113**, 1069-1074.
3. Deutsch, M. Trust, trustworthiness, and the F scale. *J. Abnorm. Soc. Psychol.,* 1960, **61**, 138-140.
4. Garfinkel, H. The rational properties of scientific and common sense activities. *Behav. Sci.,* 1960, **5**, 72-83.
5. Kelly, E. L., & Goldberg, L. R. Correlates of later performance and specialization in psychology. *Psychol. Monogr.,* 1959, **73**, (12, Whole No. 482).

6. LIEBERMAN, B. Human behavior in a strictly determined 3 × 3 matrix game. *Behav. Sci.*, 1960, **5**, 317-322.

7. LUCE, R. D., & RAIFFA, H. Games and Decisions. New York: John Wiley & Sons, 1957.

8. SCODEL, A., RATOOSH, P., & MINAS, J. S. Some personality correlates of decision making under conditions of risk. *Behav. Sci.*, 1959, **4**, 19-28.

9. VON NEUMANN, J., & MORGENSTERN, O. Theory of Games and Economic Behavior. (2nd Ed) Princeton: Princeton Univ. Press, 1947.

Department of Psychology
Ohio State University
Columbus 10, Ohio

Published as a separate and in *The Journal of Psychology*, 1961, 52, 63-75.

FORMING AND ACTING ON IMPRESSIONS*

Department of Psychology, Oberlin College

CELESTE McCOLLOUGH

A. INTRODUCTION

In his "Reanalysis of 'Impressions of Personality,'" Julius Wishner (9) has suggested a direction for further research on the problem of "how we come to know others," first studied experimentally by Asch (1). Bringing to his reanalysis the correlational techniques of semantic research and the resources of data-processing machines, Wishner is able to offer what he takes to be an operational analysis of Asch's concept of "central" and "peripheral" traits.

Between the conceptions of Wishner and Asch there is a distinction of more general interest, which Wishner has failed to detect. Asch treats the subject's behavior as involving two phases: (1) a phase of impression *formation,* in which the subject responds covertly to impression-material presented him, and (2) a phase of impression *testing,* in which he responds overtly to a set of test materials and makes use of the impression already formed. Wishner focuses his reanalysis upon the second phase alone. This difference has important consequences for the way each deals with the "central-peripheral" concept. Asch defines a trait as "central" in terms of its relationship to other traits presented with it in the material for impression formation: for him, a central trait has already performed its function before the subject is given his overt response task, and a variety of different testing conditions may be employed to demonstrate its more profound effect. Wishner, however, redefines as "central" those traits which have a certain relationship to the test-material; for him, a trait does not become "central" until the subject undertakes to respond to some particular test situation. Asch was concerned with the formation of an impression, while Wishner analyzes only the action resulting from the impression.

The genuine contrast between Asch's and Wishner's conceptions of the problem is worth pointing out because of the ease and frequency with which a covert organizing process is overlooked by psychologists concerned primarily with the task given to elicit overt response. Since the only method for

---

* Received in the Editorial Office on January 30, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

studying such a process is the observation of behavior, some doubt persists among psychologists as to whether any inference about it can be made at all. By holding constant the test-material and varying the impression-material, Asch manages to say some things about the formation process on the basis of observations of objective behavior; thus a comparison of the Asch and Wishner articles illuminates an important area in which psychological theory shows confusion.

B. ASCH'S STUDY OF "CENTRAL" AND "PERIPHERAL" TRAITS

In the first five of Asch's ten experiments, the concept of "central" and "peripheral" traits is developed. Asch read to his subjects a list of terms (e.g., "intelligent—skillful—industrious—warm—determined—practical—cautious"), with instructions to regard these terms as characteristics belonging to a particular person whom the subject was later to characterize in a few sentences. After two hearings of this impression-list (I-list), the subject carried out the following tasks: (*a*) writing of a brief sketch of the person described; (*b*) selection of the more fitting term from each pair of a test-list of paired traits, mostly opposites (T-list); (*c*) ranking of the seven I-list terms, in order of degree of influence on the impression formed; and, in the case of Experiment V, (*d*) writing down synonyms corresponding to certain I-list traits.

From these experiments, Asch concluded that some traits have greater influence than others on the final impression. Certain traits profoundly affect the cognitive content of other traits in the list, and such traits also affect the weight given to other traits in forming the impression. Those traits which "determine the content and function of other traits" (1, p. 270) he called "central"; the other traits he called "peripheral." It is worth considering the behavioral evidence which led to this conclusion:

(1) When one group was read the I-list illustrated above, and another group was read the same I-list with "cold" substituted for "warm" (Experiment I), the descriptive paragraphs written by members of the two groups differed markedly. The "warm" group offered positively-toned descriptions of sincere, well-intentioned individuals; the "cold" group generally described less attractive persons who were snobbish, calculating, unsympathetic, and the like.[1]

(2) The substitution of "cold" for "warm" produced marked group

---

[1] The effect of "warm" and "cold" in such a context has been shown also in a more naturalistic situation, where subjects responded to an actual person introduced to them in ways differing only in the word "warm" or "cold." (2, pp. 218-219)

differences in the response to ten of the 18 T-list pairs, but not in the response to the eight remaining pairs. When, instead of "cold" or "warm," the words "polite" or "blunt" were substituted in the same list (Experiment III), the T-list selections of these two groups did not differ from each other to nearly the same extent as those in Experiment I. Hence, "warm-cold" is apparently more influential or "central" than "polite-blunt." It is this part of Asch's evidence to which Wishner has confined himself.

(3) When the other six items of the I-list were read with no mention whatever of either "warm" or "cold" (Experiment II), the group responses to the T-list were intermediate between the "warm" and "cold" groups of Experiment I. However, the item "warm-cold" appeared in the T-list. When the total group was subdivided, on the basis of responses to "warm-cold," the responses of the subgroup selecting "cold" resembled those of the other "cold" group. The subjects showed the influence of their own conjectures about the "warmth" or "coldness" of the person, even in the absence of any direct information.

(4) Subjects in Experiment I often ranked "warm" or "cold" as highly influential in determining their impressions; subjects in Experiment III very rarely gave "polite" or "blunt" even second or third place in their rankings. When the other I-list traits forming its context were changed (e.g., "vain—shrewd—unscrupulous—warm—shallow—envious"), "warm" and "cold" lost heavily in the significance ascribed to them in subjects' ratings (Experiment IV).

(5) Alteration of the context of I-list traits in Experiment IV produced a transformation of the meaning of "warm" or "cold"; the same trait-name no longer represented the same trait, "warm" in the new context being interpreted as "insincere" warmth. Experiment V showed the transformation effect even more strikingly: in a context of "kind—wise—honest," subjects offered as synonyms for the I-list term "calm" such traits as "serene" and "good-natured," whereas in a context of "cruel—shrewd—unscrupulous," they offered "calculating," "icy," and the like as synonyms for the same word.

These findings give behavioral meaning to the concept of a "central" trait. Such a trait (1) sets the tone of descriptive paragraphs, (2) directs the choice of terms from the T-list, (3) affects the proposal of synonyms for other terms in its I-list context, and (4) appears to subjects to be outstandingly influential in determining their impressions. Because the effect of the "central" trait is evidenced in several different kinds of behavior, the concept does not depend for its definition on the accidental conditions of any one of the response tests alone. This is an important matter, for the concept "central" is intended to apply to the process of impression *formation,* a process which follows closely upon the presentation of the I-list items and which profoundly affects the later behavior of subjects in a variety of response tests. The centrality of a trait for the impression-formation process is dependent upon the conditions at the time of presentation of the I-list; thus, a trait possessing "centrality" in one I-list context may lose its centrality when it is grouped with other I-list traits. The implication of this finding is that an appropriate change in the I-list conditions will produce important changes in behavior, even though the task given to elicit overt response remains unchanged.

By offering this evidence for the independent influence of the impression-material on behavior, Asch stressed the role which conditions affecting impression-formation play in determining observable behavior. The remaining five experiments in his study further demonstrate this influence, through the effects of changes in instructions and in order of presentation of the I-list items; it will not be necessary to discuss these studies here, since they do not bear on the main point of Wishner's reanalysis.

C. Wishner's Reanalysis

In his reanalysis, Wishner undertakes to define "central" and "peripheral" in operational terms which will make it possible to determine in advance which traits will turn out to be central and which peripheral. He chooses to apply his analysis exclusively to one of the response measures, the selection of T-list items, and he sets out therefore to show that it is possible to predict, by correlational methods, which of the T-list pairs will be treated differently as a result of variation of one I-list item. Specifically, he reasons as follows: (*a*) "warm-cold," a central trait, produced marked group differences in response to ten of the 18 T-list pairs, while "polite-blunt," a peripheral trait, produced very slight group differences in response to all but one or two of the same pairs; (*b*) a trait will be more likely to produce marked differences in the response to items to which its own content is in some way related, than to affect those to which its content is generally unrelated; therefore, (*c*) a "central" trait can be operationally defined as one which is closely related to most of the T-items to be judged, and a "peripheral" trait as one which is unrelated to most of the T-items. It follows, further, that from an independent determination of the degree of relation between items, it should be possible to predict which items will appear central and which peripheral for subjects; and by judicious selection of a T-list containing mostly related (or mostly unrelated) items, *any* I-list item might be made to appear central (or peripheral). Other I-list items would be expected to affect the outcome only to the extent to which they, by virtue of their own correlations with T-list items, tend to exert a competing influence on the T-list choices.

A significant point concerning this analysis may be noted at once. Wishner gives no attention at all to the process of *forming* impressions. Instead, he offers an account of the experimental results wholly in terms of the conditions of the overt response test. He approaches the problem of defining a "central" trait with his attention riveted on these conditions. By Wishner's definition, a trait is "central," not because of its relation to the other I-list items, but because of the adventitious selection of the T-items used to test its effect. With such a definition, no trait is "central" or "peripheral" until it is known what response-test is to be applied. Furthermore, Wishner proposes to shift a trait's function from central to peripheral by manipulating the T-list, not by Asch's method of altering the I-list.

Late in his paper, Wishner acknowledges in a general way the role of the impression-material presented and claims that his procedure gives operational definition to this role. It is hard to understand this claim. It will be seen shortly that nowhere in his analysis does he actually give a place to the "specific implications of each term within the context or set already established by the initial instructions and by the previous terms" (9, p. 109), which he admits that we "need to know." When he mentions the correlations among the I-list items, and the correlations between the other I-list items and the T-items, he regards these merely as "suppressor" factors, obscuring the effect of the correlations on which prediction is to be based. Wishner may have intended to deal with the important influence of the impression-material, and he seems to believe he has done so; but the consideration of this material, or of the "set" formation process, is clearly absent from his reanalysis.

It is already evident that, whether or not Wishner succeeds in striking his target, the target is not quite what he took it to be. At best he may succeed in proving that the effect of a central trait shows up most clearly in the response to items to which it is most relevant, and that by using relevant or irrelevant items in the T-list one can bring out or obscure this effect. Wishner's method of determining relevancy (correlation), in obtaining his evidence, contains an interesting innovation. He had each of 214 students rate one person (his laboratory instructor—eight persons in all were rated by members of ten different groups) on all 53 trait-pairs used by Asch. Contingencies between trait-pairs were determined, indicating the extent to which any two trait-pairs tended to be used repeatedly in the same relation to each other (e.g., if a high rating was given on "warm," then "sociable" tended to show a high rating also; if "cold," then "unsociable"). The innovation lies in Wishner's use of the ratings of actual persons instead of having his subjects deal with abstract words in the manner of Bruner, Shapiro, and Tagiuri (4).[2]

With these experimentally determined correlations, Wishner first seeks to show that the differing effects of "warm-cold" and "polite-blunt" on T-list responses of Asch's subjects can be attributed to the amount of correlation between these items and the T-items. The demonstration is on the whole successful. With only two exceptions, all nine[3] of Asch's T-list traits which, in Experiment I, differentiated well between "warm" and "cold" groups have correlations of 0.30 or greater with "warm-cold," and all but one of the eight non-differentiating T-items have correlations below 0.30. If the absolute magnitude of these correlations is itself correlated with the magnitude of the differences between groups, a rho of 0.62 is obtained. Applying the same technique to Asch's Experiment III, Wishner finds the T-item correlations with "polite-blunt" somewhat lower on the whole (but not much lower, for eight items still show correlations above 0.30) and the over-all value of rho is 0.43. Wishner is understandably somewhat dissatisfied with this result, which he thinks may be rendered less striking than expected by the "suppressor" effect of correlations (1) within the I-list and (2) between other (peripheral) I-list items and the T-items.[4]

---

[2] This procedure is commendable for directing the attention of the raters to actual persons, so that the relevance of terms to the description of persons must be constantly borne in mind. It could, of course, be questioned whether the rating of eight persons, all of whom were instructors, is likely to produce correlations representative of the verbal habits used in describing "adult males" in general. But apart from the problem of selecting persons to be rated, the results of such a procedure may be different from those obtained through inferences about abstract words, and the difference may be of real interest.

[3] Wishner found it necessary to omit one of the original items ("wise-shrewd"), since the meaning seemed vague. He also used combined data from the Asch experiments and a replication of his own (6) in these comparisons.

[4] It is one of the merits of Wishner's study to have called attention to the interrelationships between items. But an examination of Wishner's data shows that correlations existing between items within the I-list will not account for the emergence of a central trait in Asch's sense, any more than correlations between

In an effort to provide a purer demonstration that centrality can be predicted from knowledge of the overt response task, Wishner is led to carry out an experiment with a newly devised I-list. In commenting on this final experiment, it is not my intent to minimize the importance of the task given to elicit overt response. Surely it is obvious that what one is able to observe depends upon the task which he sets his subject. But it is possible to choose testing conditions which effectively obscure what might otherwise be observed. This is exactly what Wishner accomplishes in his final experiment, the experiment with a "simplified" list which eliminates the "complications" of interactions within the I-list and between the non-central I-list items and the T-items. By choosing his T-list to contain six trait-pairs correlated with the central pair, "humane-ruthless," and six uncorrelated trait-pairs, none of these items being correlated with any of the remaining six I-list items ("scrupulous—cautious—good-looking—flexible—serious—fearless"), Wishner is able to show virtually perfect correspondence to his predictions. The six correlated traits show significant differences between "humane" and "ruthless" groups, while no other significant differences appear. But, in accomplishing this feat, Wishner also removes the possibility of discovering in the subject's responses any transformations in meaning which have occurred as a result of interactions within the I-list. If "scrupulous" shifts from a predominantly "ethical" meaning in the "humane" context to suggest the meticulous planning of crime in the context of "ruthless," no indication of this shift is likely to appear in Wishner's data; no word in the T-list is sufficiently related to "scrupulous" to be affected. The subject is simply not asked about his conception of the term.

D. ROLE OF THE "PERCEPTUAL MODEL"

I do not wish to appear to argue that the behavior shown by Asch's subjects is unpredictable, or that the "central-peripheral" concept is not available for operational definition. But it does seem important to point out that these tasks cannot be accomplished wholly in terms of the conditions of overt response testing. While these conditions are important, their major im-

---

this trait and the T-list will do so. Multiple correlations might work better, but they are still an indirect way of dealing with the I-list interactions. While some technique for pinning down the relations between the content of two items seems necessary, and while a correlational technique may be of value, it is to be doubted whether correlations obtained between items outside the context of their I-list will serve; it would be surprising if correlations obtained for "calm" in the context of "kind" followed the same trends as those in the context of "cruel." Wishner hopes that he has taken care of the "context" factor by his correlations; but if the context factor is taken seriously, it must be expected to alter the correlations.

portance is evidently in determining the extent to which the effects of a central trait are visible, and not in defining the role of a central trait in the formation of impressions. Whether a trait is central, in a given context, is very likely dependent upon the degree and kind of relationship it bears to other parts of the context; it may also be dependent upon the extent of its relationship to the entire population of possible trait-names which might be employed in responses. But its centrality, in Asch's sense, is not dependent upon the adventitious selection of a response-list.

But there is a matter of deeper significance here, which comes to light only when the question is asked, What might Wishner have done better, in his effort to move the study of impressions toward a more precise formulation? He limited himself to the data from T-list responses, ignoring the other measures which Asch offered as behavioral evidence of the centrality of "warm-cold" in certain contexts. Ought he perhaps to have chosen to predict one of these other measures, or even to have undertaken a prediction of *all* the types of response obtained? At once, however, it can be seen that such an undertaking would have presented no special obstacle to his method of proceeding; his technique might be used equally well to predict what words will be spontaneously chosen in writing a character sketch, or what synonyms will most frequently come to mind for a given I-list item. A prediction even of the rankings given to traits would have been an improvement only if the interaction among I-list items had been included among the determining factors. The principal fault of the reanalysis does not lie in the particular response-measure chosen, but in the attempt to do a straight-through prediction, from particular I-list item to particular response, with no pause to consider the process of impression-formation itself.

To allege this brings one up sharply against an issue which has become badly confused. Surely, on the one hand, there is clear evidence that something has happened *before* the subject is presented with the response-test. He does not wait for the T-list to be placed in front of him, before he knows that "warm" in the context of "shallow" and "unscrupulous" is not the same warmth he values so highly in other contexts. Yet it is easy, at the same time, to be misled by the fundamental rule of objective psychology into believing that *forming* an impression and *acting* upon it cannot be behaviorally distinguished, because both can be objectively studied only through the observation of behavior. It is one of the continuing contributions of Asch's study of impression formation, to have demonstrated that these processes can indeed be behaviorally distinguished.

It is, however, one of the unintended contributions of Wishner's reanalysis to have shown a danger that lies in the effort to dispense with all but those processes which can be shown to be at work in the final response-production. There is undoubtedly a strong appeal in the view that, after all, we have only the *behavior* to explain, and if we can predict the behavior successfully, we need not bother about processes presumed to lie behind it. But the examination of Wishner's final experiments with "simplified" lists ought to give one pause. It is uncomfortable to think that we might have to neglect all evidence of the presence of intervening processes, in order to purify our predictions.

Early in his article, Wishner contrasts the "perceptual model" clearly visible in Asch's discussion with a "concept formation" model, which he favors. In the concept formation model, attention centers on the task of predicting what response will occur under some particular conditions of overt response testing. In the perceptual model (which a Gestalt psychologist would certainly favor also in accounting for the formation of concepts), attention is given to the role played, in the determination of final behavior, by processes occurring when the original impression-material is presented. Actually, these terms are somewhat misleading, for they appear to refer to a difference in *content,* while the real distinction is between *modes of analysis.* The "perceptual model," which Asch called the "structural model," is not confined to problems of perception, although it does represent an attempt to study the relevance of perceptual laws, particularly the laws of perceptual organization, to other psychological problems. Neither does the "concept formation" model apply only to conceptual problems. Furthermore, the word "formation" does not refer, in fact, to any "formation" process, but is simply part of the technical name for a certain kind of psychological experiment.

A disposition to reject the "perceptual" mode of analysis as weak and elusive appears also in studies of other sorts of cognitive behavior. After a careful experimental and logical analysis of Köhler's theory of association, Postman and Riley (7) conclude:

> The Gestalt theorist is concerned with the effects of perception on memory . . . . In the analysis of learning and retention the Gestalt theorist is thus forced to maintain a precarious distinction between an initial experience (perception) and its later consequences (memory). It is difficult to specify where one ends and the other begins . . . . Yet, the Gestaltist insists on looking for the influences of primary perceptions on memory. In his preoccupation with the persisting effects of

> phenomenal experience, he pays little attention, if any, to the conditions which influence *S*'s responses from which inferences about both perception and memory must be made. (7, pp. 71-72)[5]

The probability, however, is that the distinction emphasized by the "perceptual" model not only can but must be made. Postman and Riley make this appear most unlikely by the use of the words "phenomenal experience," a phrase which suggests something inaccessible and inherently private. Prentice has pointed out (8, p. 437) that Gestalt psychologists "have been willing to make inferences about the phenomenal experience of their observers and try to write laws relating it to the independent variables of stimulation." He goes on to emphasize a crucial point: that these inferences are made from overt behavior, not from introspection. The "phenomenal experience" which the Gestaltists mean by perception is a process objectively evidenced by behavior, yet which occurs close upon the original stimulus presentation and depends largely upon those stimulus conditions. It is common for psychologists to avoid touching on such prior, organizational processes by invoking the concept of "set." Postman and Riley themselves appeal to the "set" concept (see "initial response bias," 7, pp. 66ff), and Wishner's allusion to the influence of "context or set" (9, p. 109) has already been mentioned above. It is fair to say that the frequent invoking of "set" implicitly acknowledges its role in determining behavior, but that few have tried as hard as the Gestalt psychologists to say something about set-*formation*. Indeed, as the papers of Wishner and of Postman and Riley reveal, sometimes the problem of formation is even regarded as dispensed with, once the concept of "set" has been invoked.

But the perceptual or structural model is let alone by many psychologists not only because it appears to fly in the face of objectivism, but also because it appears vague. The distinction between an elementaristic point of view (Asch's Proposition I, 1, p. 259) and a structural point of view (Proposition II, p. 259 and p. 286) is still not generally understood, for Wishner was able to mistake his approach as "implementing" the structural viewpoint. Indeed, Wishner's conception is not the raw "impression $=$ a $+$ b $+$ c $+$ d $+$ e" of a purely summative theory. It is more like a variant of Proposition I mentioned by Asch as Proposition Ib (p. 286), where the impression is the sum of the elements plus, also, separate interactions between pairs of the elements. Perhaps the misunderstanding could be remedied, if it were kept in mind that the Gestalt psychologist is trying to say something about the *organizing* process, preceding emergence of the overt response. A procedure such as that of Wishner does not touch on the interactions governing impression-formation as a process of organization. Use of the structural model, then, implies two things: (1) that attention will be given to the organizing phase, the phase of impression-formation; and (2) that stimulus elements will not be treated as if their properties are wholly independent of the context in which they appear. If these points are taken seriously, one need not fear that the structural model will require him to abandon analysis and look at everything at once; nothing need deter him from the use of analytic and quantitative procedures as a way of developing his account of behavior.

---

[5] Asch has addressed himself to the problem of the "perceptual conditions of association" in a recent monograph with Ceraso and Heimer (3).

Finally, a specific illustration of what has been meant here by the introduction of an organizing process in order to arrive at a prediction may be obtained through brief consideration of a miniature theory now available in a substantially complete form, Helson's theory of adaptation level (5). The theoretical problem is to predict the judgments which will be made when a series of stimuli (weights, colors, etc.) are presented and each must be assigned a judgment denoting its position within the series. According to an early, informal perceptual model, it was maintained that the individual stimuli would be judged within a "frame of reference" growing out of the subject's perception of the series as a whole. Helson's theory uses S-elements, combined by a geometrical averaging procedure, to predict, first, the neutral "reference point" (adaptation level or AL) resulting from their interaction. For the absolute judgment of weights, the equation is

$$AL = \sqrt[n]{S_1 \cdot S_2 \cdot \cdots \cdot S_n} + k_1.$$

The theory then proceeds by a further equation to predict the response, R_j, to a particular stimulus weight, S_j, where C represents a judgment of "medium" (5 on a 9-point scale):

$$R_j = C + k_2 \log \frac{S_j}{AL}.$$

This theory is remarkably successful in predicting responses, but it is doubtful whether it could have been so successful had it not come to grips with the preparatory process of adaptation level formation in advance of response-production. Obviously, there is nothing in the equation which cannot be described in terms of S and R; the equation for R_j could equally well have been written

$$R_j = C + k_2 \log \frac{S_j}{\sqrt[n]{S_1 \cdot S_2 \cdot \cdots \cdot S_n} + k_1}.$$

The essential quality of the equation, for our purposes, is that it contains an expression which describes the interaction of the series stimuli with each other, and which gives this organizational process its proper role in the production of the overt response.

E. SUMMARY

A comparison of Wishner's reanalysis and the original Asch studies on forming impressions of personality brings out some points about the specific problem and certain more general questions.

1. While Asch offered behavioral evidence bearing on the process of impression *formation,* Wishner has concentrated his attention on the behavior of subjects *acting* on their impressions.

2. According to Asch, a trait like "warm" or "cold" plays a central or peripheral role in a context of trait-names, depending on the other traits in the context. Wishner has redefined "central" and "peripheral" in terms of the amount of relation between a trait given as part of the original impression-material and the trait-list employed in measuring responses. Asch is concerned largely with conditions existing during the initial impression-forming phase, Wishner with those existing at the time of overt response testing.

3. Wishner's effort to show that these correlations with the response list are wholly able to account for the data from Asch's and his own earlier experiments meets with only moderate success. He therefore constructs new "simplified" lists in which indeed no influence attributable to stimulus interactions within the original impression-list is detectable, because his procedure has carefully eliminated—not the influence—but any possibility that evidence of influence might arise. It is suggested, therefore, that what Wishner has succeeded in specifying is the extent to which the effect of central traits will be *visible* in the behavioral observations.

4. The disposition to attempt response prediction directly, without attention to the impression-forming process, appears to arise from a belief that the processes of forming and acting on impressions are not behaviorally distinguishable because both can be studied only through actual behavior.

Yet Asch has been able to give behavioral reference to this distinction. Since doubt persists regarding the distinguishability of any such organizational processes, the basic divergence between Asch and Wishner takes on a wider significance. Although some writers, notably Postman and Riley, explicitly reject the consideration of organizing processes, some attention to the contextual interaction between elements during the covert organizing process is likely to be necessary in successful prediction of responses.

REFERENCES

1. Asch, S. E. Forming impressions of personality. *J. Abnorm. Soc. Psychol.*, 1946, **41**, 258-290.
2. ———. Social Psychology. New York: Prentice-Hall, 1952.
3. Asch, S. E., Ceraso, J., & Heimer, W. Perceptual conditions of association. *Psychol. Monogr.*, 1960, **74**, No. 3 (whole No. 490).
4. Bruner, J. S., Shapiro, D., & Tagiuri, R. The meaning of traits in isolation and in combination. *In* R. Tagiuri and L. Petrullo (*Eds.*), *Person Perception and Interpersonal Behavior*. Stanford, Calif.: Stanford Univ. Press, 1958. Ch. 18.
5. Helson, H. Adaptation level theory. *In* S. Koch (*Ed.*), *Psychology: A Study of a Science*, Study I, Volume I. New York: McGraw-Hill, 1959. Pp. 565-621.
6. Mensh, I. T., & Wishner, J. Asch on "Forming impressions of personality": further evidence. *J. Pers.*, 1947, **16**, 188-191.
7. Postman, L., & Riley, D. A. A critique of Köhler's theory of association. *Psychol. Rev.*, 1957, **64**, 61-72.
8. Prentice, W. C. H. The systematic psychology of Wolfgang Köhler. *In* S. Koch (*Ed.*), *Psychology: A Study of a Science*, Study I, Volume I. New York: McGraw-Hill, 1959. Pp. 427-455.
9. Wishner, J. Reanalysis of "Impressions of personality." *Psychol. Rev.*, 1960, **67**, 96-112.

Department of Psychology
Oberlin College
Oberlin, Ohio

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 77-86.

TASK DIFFICULTY AND ANXIETY LEVEL AS FACTORS AFFECTING PERFORMANCE IN A VERBAL LEARNING SITUATION*[1]

Department of Psychology, Tufts University

BERNARD W. HARLESTON AND SCOTT M. CUNNINGHAM

Within the framework of a paired-associates learning situation the present research focused upon (1) an analysis of the relative effects of meaningfulness of stimulus and response members in determining difficulty level, and (2) an analysis of the relationship between anxiety level and meaningfulness of stimulus and response members in determining performance.

A series of studies by Noble *et al.* (Noble and McNeely, 1957; Cieutat, Stockwell and Noble, 1958) has shown that in a paired-associates learning task, the meaningfulness of the stimulus items in relation to the meaningfulness of their paired response items is a reliable and significant factor affecting the difficulty of the task. Among four lists in which the meaningfulness of the stimulus items in relation to the meaningfulness of the response items was carefully varied, Noble et al. found that the High-High list, consisting of stimulus items and response items of high *m*-scale values, was least difficult to learn, followed in order of increasing difficulty by a Low-High list, a High-Low list, and finally, a Low-Low list. This ordering for difficulty, along with the magnitude of the difference between lists, led them to conclude that while both stimulus meaningfulness and response meaningfulness were significant determinants of task difficulty, response factors were of greater significance. Because the present investigators sought to relate the effect of anxiety level to task difficulty level, this approach was used. However, since other investigators had not found results consistent with those of Noble and his coworkers (Sheffield, 1946; Mandler and Campbell, 1957) one purpose of the present experiment was to determine whether this order of difficulty could be reproduced with statistical reliability.

As noted earlier, the other purpose was to compare the performance of high and low anxious *S*s on each of the four list conditions studied. While

---

*Received in the Editorial Office on February 13, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Publication of this article was supported in part by a grant from the Tufts University Faculty Research Fund. Portions of this paper were presented at the 1960 meeting of the Eastern Psychological Association.

a variety of predictions might have been made, the authors sought specifically to explore the hypothesis that as the difficulty level of the task increases, the difference in performance between low-anxious (LA) and high-anxious (HA) *S*s increases, with the high-anxious *S*s performing increasingly more poorly than the low-anxious *S*s. A number of investigators, particularly Saltz and Hoehn (1957) and Sarason (1958) have presented data which lend some support to this hypothesis. It was felt that comparing performance of HA and LA *S*s on four lists varying in difficulty level would provide a sensitive test of this hypothesis.

A. METHOD

1. *Subjects*

Forty-eight males and 48 females served as subjects. Half of the males and half of the females had scores of eight or less (LA) on the 50-item form of the Taylor Scale of Manifest Anxiety; the remaining 24 males and 24 females had anxiety scale scores of 23 or above (HA).

2. *Learning material*

Four lists of 12 paired associate nonsense syllables were developed. List I, the H-H list, was composed of stimulus-response pairs, the members of which had an average association value of 90 per cent as reported by Glaze (Hilgard, 1951). List II, the L-H list, was composed of stimuli with an average association value of 8 per cent, and of responses with an average association value of 90 per cent. List III, the H-L list, was composed of stimuli with an average association value of 90 per cent, and of responses with an average association value of 8 per cent. List IV, the L-L list, was composed of pairs, the members of which had an average association value of 8 per cent.

Two forms of the H-H and L-L lists were developed. The forms differed in that the items serving as stimuli for some *S*s served as responses for other *S*s, and vice versa.

3. *Experimental design and procedure*

High and low-anxious males and females were assigned on a random basis to one of the four lists. Thus, six HA males, six HA females, six LA males and six LA females learned each list. All *S*s were run individually.

A modified Gerbrands memory drum was used, exposing the stimulus items for 1.5 seconds and the S-R pair for 1.5 seconds. The inter-trial interval was 4.5 seconds.

The instructions read to *S*s were standard ones for a learning task of this sort. After the first trial, during which *S* pronounced the stimuli and S-R pairs, *S* was presented the list for 50 trials, or until he was able to anticipate correctly all response items on two successive trials, whichever occurred first. During the learning trials, four different orders of the list were used to minimize serial effects.

Performance on the experimental list was preceded by eight trials on a practice list of 12 pairs of meaningful, five-letter words. The conditions of presentations were essentially the same as for the experimental lists. The interval between the end of the practice list and the beginning of the experimental list was three minutes.

B. RESULTS

Analysis of the practice data revealed that the female *S*s performed significantly better than the males ($F = 6.03$; with 1 and 80 *df*, $P = .02$). It was decided therefore, to complete separate analyses of the experimental list data for the two sexes.

The mean total number of correct responses made during the first 30 trials[2] by the HA and LA *S*s assigned to the various lists is presented in Table 1. Across anxiety levels, the order of difficulty of the lists for both males and females was as predicted by Noble et al.

TABLE 1
Mean Number of Correct Responses, and Standard Deviations, for Thirty Acquisition Trials for Low and High Anxious Females and Males on the Various Lists

	Females				Males			
	LA		HA		LA		HA	
List	MN	SD	MN	SD	MN	SD	MN	SD
H-H	192.17	75.51	186.50	81.03	174.67	62.04	167.30	87.84
L-H	171.67	58.76	149.38	70.23	182.00	69.74	127.00	87.53
H-L	121.17	22.45	139.50	58.45	132.50	89.21	127.67	45.18
L-L	123.30	73.23	79.88	39.39	98.30	36.70	101.00	45.86

There are, however, irregularities when HA and LA groups are viewed separately. The learning functions for HA and LA *S*s on the various lists are presented graphically in Figures 1a and 1b, based upon trial blocks of five trials.

---

[2] Although the criterion was 50 practice trials or two errorless trials, whichever occurred first, it was decided to use the data of the first 30 practice trials for statistical analysis because of the extensive variability between lists. In actual fact, the same order and direction of results are found with either the 30 trial or the 50 trial data.

In tabulating the data for those *S*s who reached the two trial criterion before 30 trials, the assumption was made that any responses on trials beyond the criterion would have been correct and were so recorded.

Separate analyses of variance were done for the females and males. The design used in each case was a 2 × 2 × 2 × 6 repeated measures design, with trials grouped in blocks of five. The summaries are presented in Table 2. For both females and males, the only statistically significant between-group effect was that due to variations in meaningfulness of the response items. When the response members were of high association value, learning

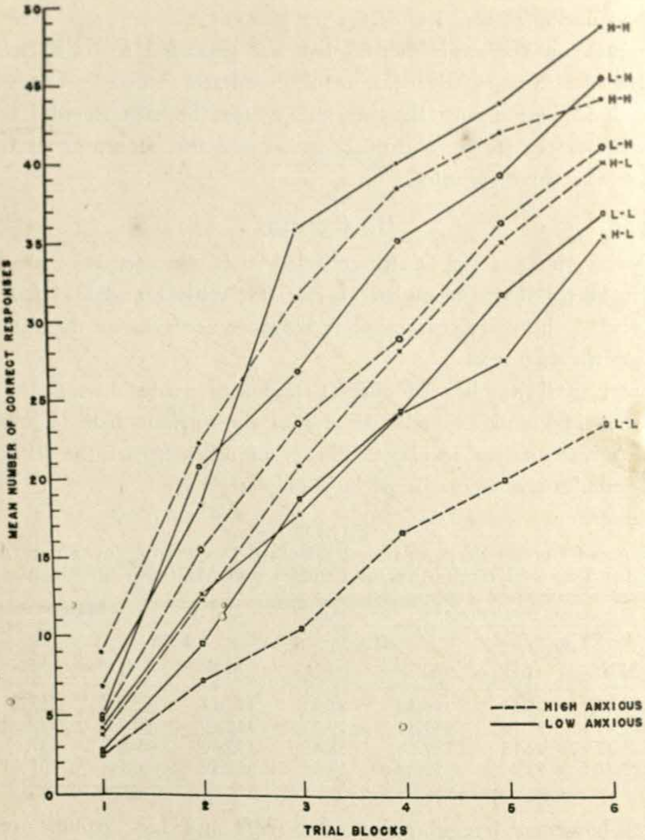


FIGURE 1a
MEAN NUMBER OF CORRECT RESPONSES IN EACH OF SIX BLOCKS OF FIVE TRIALS BY LOW AND HIGH ANXIOUS FEMALES ON THE VARIOUS LISTS

was significantly faster. Subsequent analysis by *t*-tests (Lindquist, 1953) indicated that for the females the differences due to the association value of the response items were significant at both levels of meaningfulness of stimulus items; for the males, the differences due to the association value of the response items were statistically significant only when the stimulus items were of low association value. Meaningfulness of stimulus items, anxiety level and all between-group interactions failed to reach statistical significance.

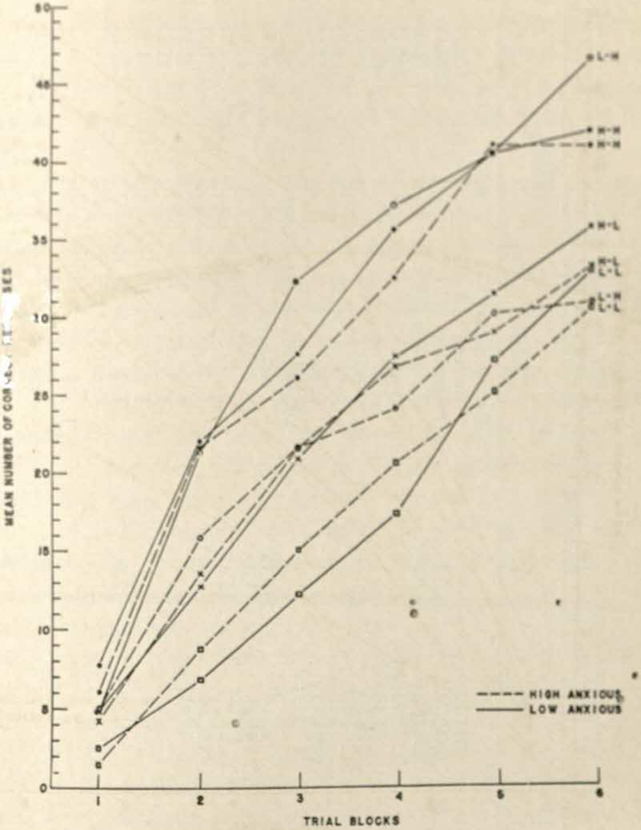


FIGURE 1b
MEAN NUMBER OF CORRECT RESPONSES IN EACH OF SIX BLOCKS OF FIVE TRIALS BY LOW AND HIGH ANXIOUS MALES ON THE VARIOUS LISTS

The within group analyses based upon trial blocks produced, for the

TABLE 2

Summary of Analyses of Variance of Total Number of Correct Responses in Thirty Acquisition Trials (Grouped in Blocks of Five) by Low and High Anxious Females and Males

Females				Males			
Source	*df*	*ms*	*F*	Source	*df*	*ms*	*F*
Between *S*s	47			Between *S*s	47		
Anxiety (A)	1	351.12		Anxiety (A)	1	522.73	
Stimulus (S)	1	1662.72	2.15	Stimulus (S)	1	1096.69	1.17
Response (R)	1	6962.00	9.0**	Response (R)	1	4576.06	4.88*
A × S	1	767.02		A × S	1	203.33	
A × R	1	1.13		A × R	1	499.99	
S × R	1	0.00		S × R	1	95.67	
A × S × R	1	253.12		A × S × R	1	378.15	
Error (b)	40	773.42		Error (b)	40	937.59	
Within *S*s	240			Within *S*s	240		
Trials (T)	5	7767.10	236.37***	Trials (T)	5	6706.50	194.79***
T × A	5	33.60	1.02	T × A	5	38.06	1.11
T × S	5	38.32	1.17	T × S	5	20.57	
T × R	5	152.80	4.65***	T × R	5	95.91	2.79*
T × A × S	5	30.94		T × A × S	5	22.38	
T × A × R	5	4.40		T × A × R	5	35.63	1.03
T × S × R	5	25.69		T × S × R	5	33.29	
T × A × S × R	5	84.20	2.56*	T × A × S × R	5	30.03	
Error (w)	200	32.86		Error (w)	200	34.43	
Total	287			Total	287		

* $p < .05$
** $p < .01$
*** $p < .001$

females, a significant interaction between trial blocks and meaningfulness of response items and a significant trials by anxiety by stimulus by response interaction. The former significant interaction means that the superiority of performance on lists with the more meaningful response items generally increased as the trials progressed. The significant TXAXSXR interaction is due to the increasingly superior performance of LA over HA *S*s on the L-L list and the less dramatic but still increasingly superior performance of HA over LA *S*s on the H-L list. For the males, only the interaction of trial blocks and meaningfulness of response items was found to be statistically significant.

As noted earlier, no statistically significant between-group difference was found for anxiety level, either as a main effect or in interaction. The hypothesis under examination was that as task difficulty increased the differences in performance between HA and LA *S*s would increase. It had been expected, of course, that the lists would reliably order themselves in terms of difficulty, thereby allowing an adequate test of *this* hypothesis over a wide range of difficulty levels. The failure to find a completely reliable ordering makes the testing of this hypothesis more difficult. If one compares the differences in performance between HA and LA *S*s on the H-H list and the L-L list, it is observed that the differences do become greater for the females; for the males, the magnitude of differencs becomes less (and falls in the opposite direction). Perhaps the most significant aspect of these data is that the list producing the greatest differences in performance between high and low anxious *S*s differed for the females and the males. The list producing the greatest difference between high and low anxious females was the L-L list, the list ranking as quite difficult for all *S*s. On this list the difference in favor of the LA *S*s, was of the magnitude of a mean of 43.42 correct responses. No other list produced as great a mean difference (H-H: 5.67; L-H: 22.34; H-L: —18.33). The list producing the greatest difference between high and low anxious males was the L-H list, which produced a mean difference of 55, again in favor of the LA *S*s. The other mean differences were: H-H: 7.37; H-L: 4.83; L-L: —2.70. However, when these conditions of maximum differences between high and low anxious *S*s were subjected to tests of statistical significance based upon the *t*-tests, the *t*-ratios failed to reach statistical significance.

C. DISCUSSION

Considering first the relationship between meaningfulness of stimulus and response item and difficulty of the task, the data reported here lend only partial support to the statistical reliability of the order of difficulty reported by Noble. While a clear superiority was found for high meaningful responses over low meaningful responses, the superiority of high meaningful stimuli over low meaningful stimuli did not emerge as clearly. These results, when viewed along with those of Sheffield (1946), Mandler and Campbell (1957), and the recently published work of Hunt (1959), indicate that attempts to vary difficulty level by manipulating the meaningfulness of the stimulus items are certainly less likely to produce statistically reliable differences than attempts based upon manipulating the meaningfulness of the response items. The failure to support definitely the Noble formulation may be due to any of several procedural differences which have characterized much of the research in this area. Among such procedural differences would be included the kind of verbal material used, the rates of presentation during the learning trials, the exposure durations, the extent and kind of warmup activity, and the acquisition criteria. While these differences in procedure make difficult any direct comparisons among studies, the fact that performance *may* be sensitive to such differences suggests that the order of difficulty found by Noble may be more restricted than originally assumed.

No support was found for the hypothesis that the differences in performance between high and low anxious *S*s would increase as the task became more difficult. It was found, however, that for the females, the greatest difference between HA and LA *S*s occurred on a very difficult list, the L-L list, while the L-H list produced the greatest difference between HA and LA males. A provocative hypothesis suggested by these findings is that the effects of anxiety level depend, in part, upon the relationship between the difficulty level of the task and the subjects' ability. As the difficulty level of the task approaches some critical relationship to ability level, the interfering effects of high anxiety level increase; as the difficulty level of the task extends beyond the critical relationship to ability level, the interfering effects of high anxiety level are sharply reduced. According to this formulation we would say, then, that for the females, with their better ability as reflected on the practice list, the L-L list did not exceed the critical relation between difficulty and ability; thus, it was most sensitive to differential anxiety effects. For the males, however, the more difficult lists (H-L and L-L) exceeded the critical relation between difficulty level and ability. Performance on these were thus less sensitive to anxiety effects.

At present the hypothesis is admittedly *ad hoc* and quite generally formulated. Research is currently in progress in our laboratory aimed at collecting data in support of the hypothesis and testing certain implications of this formulation. Independent of the validity of this formulation, the data reported here bearing on the role of anxiety clearly demonstrate, as have those from many other studies, that if anxiety level is a determinant of performance in verbal learning situations, it is a variable highly sensitive to the influence of other variables.

After the present research was in progress, a paper was published by L'Abate (1959) in which an approach similar to that reported here was followed. Predicting that HA *S*s would perform *better* than LA *S*s on lists with low association value, because there were no competing responses, L'Abate had HA and LA *S*s learn one of four lists (H-H, L-H, H-L, L-L) of nine nonsense syllable pairs. He found a significant *F* for lists, but no reliable effects due to anxiety level. Although his study was similar in certain respects to the present one, the markedly uneven distribution of *S*s over his four lists, certain procedural and statistical variations, and lack of sufficient descriptive information, make it both difficult and somewhat inappropriate to compare the two studies directly.

D. SUMMARY

Independent groups of LA and HA females and males learned one of four lists of paired associate nonsense syllables. The lists differed in terms of the relationship of the association value of the stimulus items to the response items. It was found that increasing the association value of the response items facilitated learning; variations in the association value of the stimulus items failed to affect learning.

No statistically significant main effects due to anxiety level were found. However, it was noted that the list producing the greatest difference between HA and LA *S*s differed for the females and the males. For the former, the low-low list, consisting of stimuli and responses of low association value produced the greatest difference. The difference between LA and HA males, again in favor of the LA *S*s, was greatest on the low-high list, the list with stimuli of low association value and responses of high association value.

On the basis of these data and the observation that the females learned the practice list significantly faster than the males, a tentative hypothesis was formulated which proposes that anxiety effects depend upon both task difficulty level and an ability factor.

REFERENCES

1. CIEUTAT, V. J., STOCKWELL, F. E., & NOBLE, C. E. The interaction of ability and amount of practice with stimulus and response meaningfulness (*m*, *m'*) in paired associate learning. *J. Exp. Psychol.*, 1958, **56**, 193-202.
2. HILGARD, E. Methods and procedures in the study of learning. *In* S. S. Stevens (*Ed.*), *Handbook of Experimental Psychology*. New York: Wiley, 1951.
3. HUNT, R. G. Meaningfulness and articulation of stimulus and response in paired-associate learning and stimulus recall. *J. Exp. Psychol.*, 1959, **57**, 262-267.
4. L'ABATE, L. Manifest anxiety and the learning of syllables with different associative values. *Amer. J. Psychol.*, 1959, **72**, 107-110.
5. LINDQUIST, E. F. Design and analysis of experiments in psychology and education. New York: Houghton Mifflin, 1953.
6. MANDLER, G., & CAMPBELL, E. H. Effect of variation in associative frequency of stimulus and response members on paired associate learning. *J. Exp. Psychol.*, 1957, **54**, 269-273.
7. NOBLE, C. E., & MCNEELY, D. A. The role of meaningfulness (*m*) in paired associate learning. *J. Exp. Psychol.*, 1957, **53**, 16-22.
8. SALTZ, E., & HOEHN, A. J. A test of the Taylor-Spence theory of anxiety. *J. Abnorm. Soc. Psychol.*, 1957, **54**, 114-117.
9. SARASON, I. G. Effects on verbal learning of anxiety, reassurance, and meaningfulness of material. *J. Exp. Psychol.*, 1958, **56**, 472-477.
10. SHEFFIELD, F. D. The role of meaningfulness of stimulus and response in verbal learning. Unpublished doctoral dissertation, Yale Univ., 1946.

Department of Psychology
Tufts University
Medford 55, Massachusetts

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 87-97.

THE DISCONTINUITY IN SCHMEIDLER'S ESP-RORSCHACH DATA*[1]

Department of Biophysics, University of Pittsburgh

R. A. McConnell

In a recent monograph G. R. Schmeidler (2) presented Rorschach and ESP data from more than 1000 subjects gathered over the period 1945-1951. The present paper will examine certain aspects of the data of Table 10 of that monograph.

Table 10 is concerned with the first of the two test hypotheses of the monograph, namely (Chapter III) one relating ESP performance to social adjustment—the latter as measured by the Check List score from Ruth Munroe's Rorschach Inspection Technique. The other hypothesis (Chapter IV) supposes a relation between ESP scores and certain *ad hoc* "signs" derived from the same scoring of the Rorschach protocols. Thus, as Schmeidler indicates (p. 26) these adjustment and sign hypotheses are in certain ways inseparably related.

Figure 1 below, taken from the sixth column of Table 10, is a histogram of the total deviation of ESP scores from chance expectation for each Munroe Check List score. The data have been separated according to the two classes of subjects, namely, the "sheep," who accepted the possibility of ESP, and the "goats," who denied that possibility.

A salient feature of Figure 1 is the discontinuity appearing between the Check List scores of 10 and 11 for both sheep and goats. With regard to this discontinuity Schmeidler (*ibid.* p. 26) makes three points, 1. these breaks in the graphs were not anticipated, 2. the line between 10 and 11 where the breaks occur is arbitrary as far as the Munroe technique is concerned, and 3. this 10-11 line was taken as the defining point for "good" and "poor" adjustment in the original plan of the ESP experiment.

Schmeidler did not present a quantitative assessment of the chance probability of this discontinuity in her monograph. That is done in Table 1 by testing for interaction in a 2 × 2 array of the number of subjects above and

---

* Received in the Editorial Office on February 13, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This paper is indexed as Publication No. 86 in the Biophysics Department of the University of Pittsburgh, and was prepared under a research program financially supported by the A. W. Mellon Educational and Charitable Trust.

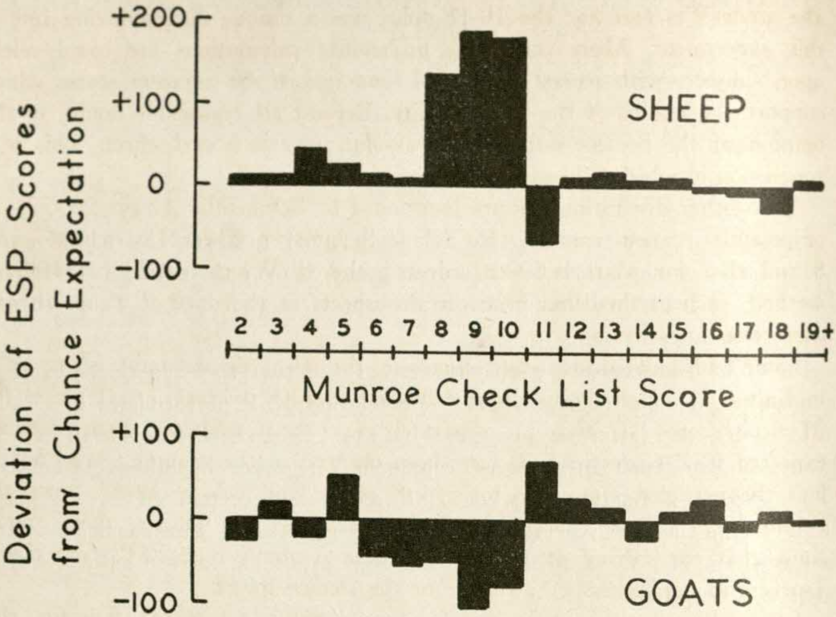


FIGURE 1
Total deviation of ESP scores for sheep and for goats as a function of Munroe's Rorschach "Check List" measure of social adjustment. (From Table 10 of Schmeidler's monograph)

below chance expectation against Check List categories 10 and 11. The chance probabilities for sheep and goats respectively are P = .0004 and P = .003. It might be argued that these values need correction for selection of the Check List 10-11 score point. In reply, it should be said that although

TABLE 1
INDEPENDENCE OF ATTRIBUTES AMONG SHEEP AND GOATS WHO HAD MUNROE CHECK LIST SCORES OF 10 AND 11 AND ESP SCORES ABOVE AND BELOW CHANCE EXPECTATION

ESP Scores	Sheep Check List Scores 10	11	Totals	ESP Scores	Goats Check List Scores 10	11	Totals
Above Chance	42	20	62	Above Chance	13	28	41
Below Chance	18	36	54	Below Chance	27	13	40
Totals	60	56	116	Totals	40	41	81

$\chi^2 = 12.34$ d.f. = 1
P = .0004

$\chi^2 = 9.00$ d.f. = 1
P = .003

the analysis is *post hoc* the 10-11 point was a unique dichotomizing line in this experiment. Moreover, these probability calculations are based solely upon subjects with scores 10 and 11, and ignore the adjacent scores which support the reality of the discontinuity. Beyond all reasonable doubt, in the opinion of the present author, this discontinuity is a real effect. This is a conclusion in which Schmeidler presumably concurs.[2]

Two other discontinuities are mentioned by Schmeidler (2 pp. 26, 32) as of possible interest, namely, that for sheep between Check List scores 7 and 8 and that for goats between scores 5 and 6. When tested by the above method, each of these has a chance probability in excess of P = .1 without correction for selection.

Some additional information regarding the 10-11 discontinuity is presented in Figure 2, which shows the distribution of ESP average (run) scores for Munroe scores 10 and 11, separately for sheep and for goats. Chance expected ESP scores of 5.00 have been omitted in the grouping so as not to bias the expected symmetry for groups above and below chance. Since the actual mean chance expected score is around 43 per subject, these graphs show that the scoring mode for these data is about four ESP score points (successful card guesses) away from the chance mean.

Schmeidler suggests two possible explanations (2 p. 28 ff.) for this discontinuity. The first of these is that "this abrupt change in the dependent variable may be characteristic of results obtained with the Check List technique." *I.e.,* "perhaps there is a critical area where even a slight increase in the number of [Check List] deviations from the norm will swing the balance [so that] a person's feelings about himself, once the critical point has been passed, may be that he is different from his group rather than that he belongs in it."

What Schmeidler is saying is that meaningful (*i.e.,* non-chance) abrupt changes in a dependent variable can occur for a single score interval of the

---

[2] The data of the Schmeidler monograph (*ibid.*) were originally scheduled for publication in the book *ESP and Personality Patterns* (4) of which RAM is the junior author. In accordance with the plan of the experiment, the ESP scores had been dichotomized in three ways (sheep vs. goat, well vs. poorly adjusted, and with or without certain Rorschach "signs"). Within the limitations imposed by their non-orthogonality, the ESP scores as the dependent variable were given the equivalent of a 2 × 2 × 2 factorial analysis of variance, with results that were interesting and statistically significant.

The actual Munroe Check List scores upon which the social adjustment dichotomy was based were not received by RAM until the book manuscript had received final editing and had been sent to the printer. When a statistical discontinuity was discovered in those scores, the book was withdrawn from publication. It was released one year later after the deletion of three chapters dealing with the Rorschach data and with an explanation (4 p. 66 f.) of that deletion.

Munroe Rorschach Check List. Since this possibility is proposed explicitly at length (p. 28) as well as implicitly (pp. 10, 15), it deserves detailed examination.

Elsewhere (1) RAM has dealt with this problem generally, showing that a test score must be considered as an inherently continuous variable if it consists of the sum of checks received on a set of non-ordered test items and showing further that a complex functional relationship cannot be exhibited between two variables unless both variables can be measured with sufficient precision.

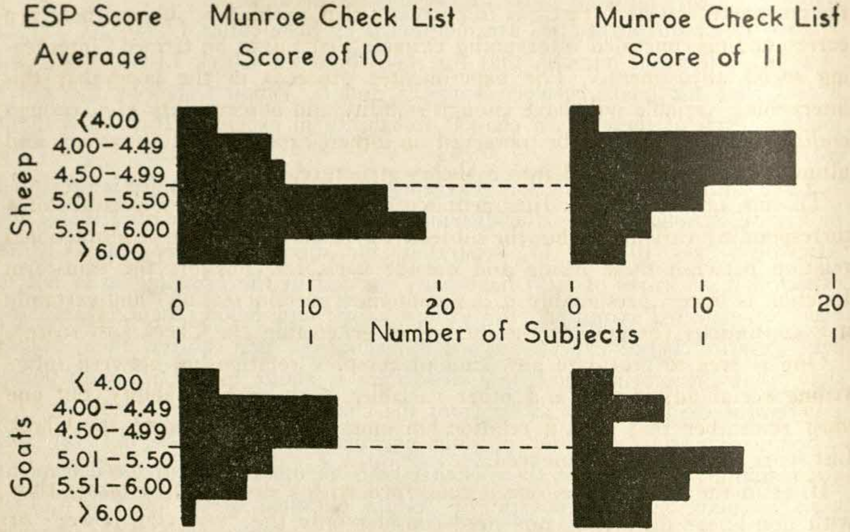


FIGURE 2
Distribution of average ESP scores for sheep and for goats having Munroe Check List scores of 10 and 11. (Subjects from Table 10 of Schmeidler's monograph, omitting those at chance expectation)

These ideas will now be applied to Schmeidler's first explanation of her Rorschach discontinuity. For her case the "complex functional relationship" approximates a step-function occurring between two successive values of the Munroe Check List score and the precision in question is the uncertainty in that score. In order to assess that uncertainty it is necessary to consider what it is that the Check List score measures.

A large number of verbal Rorschach responses are categorized and counted. The resulting Check List score is an observed variable which is tagged "social adjustment score." It is hoped that this observable can be related to other observables such as ESP performance or college grades. Moreover, the relationships that will be sought are to be of a psychological nature, *i.e.,* they will "take place inside" the subject.

But social adjustment score is necessarily an "outside" variable. So that this highly specific outside variable can lead to useful scientific generalization, the customary research tactic is to postulate that within the subject there is a corresponding concealed intervening variable that might be termed "intervening social adjustment." The experimenter proceeds in the hope that this intervening variable will have enough stability and observability (*i.e.,* enough *reality*) so that it can be observed in other experimental situations and ultimately be incorporated into a theory structure.

To sum up, the social adjustment score is conceived to be a measure of a corresponding variable within the subject. There must therefore be a functional relation between these inside and outside variables. Possibly the transform function is linear, presumably it is monotonic (ever increasing) and certainly it is continuous (because of the method of generating the Check List score).

One is free to postulate any kind of complex relationship between intervening social adjustment and other variables, such as ESP ability, but one must remember that such a relationship must be transformed to the Check List score in order to be observed.

If, as in the present case, one is concerned with a discontinuity rather than with non-linear distortion, one need consider only the "resolving power" of the transforming function, *i.e.,* the uncertainty in the Check List score resulting from the transformation of a single point on the intervening social adjustment scale. And for all practical purposes one can approximate the ability of the transforming function to transmit a discontinuity by assuming that the transforming uncertainties are normally distributed with a given standard deviation.

Uncertainties in deriving the Munroe Check List score arise from many sources such as transforming differences between subjects (even if all were supposed to have a discontinuous social adjustment at exactly the same point on the intervening scale!), the day-to-day differences of a particular subject (as they might affect his responses to the Rorschach inkblots), differences between experimenters in administering the Rorschach and in evaluating the protocols, and the uncertainty of a particular experimenter in evaluating any particular protocol.

Schmeidler estimates (*vide infra*) the last of these uncertainties alone as ranging from two to five Check List score points. The present author estimates that the transformation error from all sources combined might have a standard deviation as high as ten points (making the Check List score all but useless). Presumably no one would estimate the standard deviation of the error as less than two score points.

Schmeidler's discontinuity can be approximated by a step function as shown in Figure 3a. When such a function is given a linear transformation with a standard deviation of error equal to two score points, the result is as shown in Figure 3b. The implications of Figure 3 for Schmeidler's discontinuity are contained in the observation that the biggest step in the transformed function is no more than 20 per cent larger than the steps immediately adjacent to it and has their same direction.

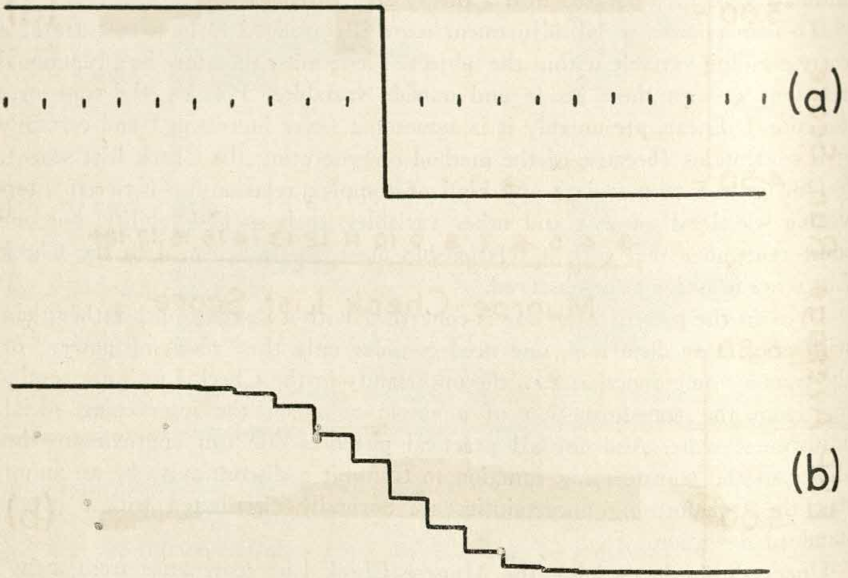


FIGURE 3
Histogram with a single step (a) and its transform (b) when the transforming function has x-axis errors normally distributed with a standard deviation of two x-axis intervals

What such a transformation means to Schmeidler's discontinuity can be made more intuitively evident by imagining that her average ESP scores (Figure 4a) did not in fact occur as a function of the observable "social adjustment score" but instead represent a real (albeit curiously complex) relation between ESP ability and the non-observable "intervening social adjustment." When transformed linearly with a standard deviation of the error of 2, the resulting ESP scores (Figure 4b) show no appreciable discontinuity. Under these imagined circumstances the observable data obtained in the experiment would show the relationship of Figure 4b plus superimposed chance fluctuations of the ESP scores.

The foregoing would appear to establish the proposition that the discontinuity found in Schmeidler's Rorschach data cannot represent a psychological relationship between social adjustment measured by the Munroe

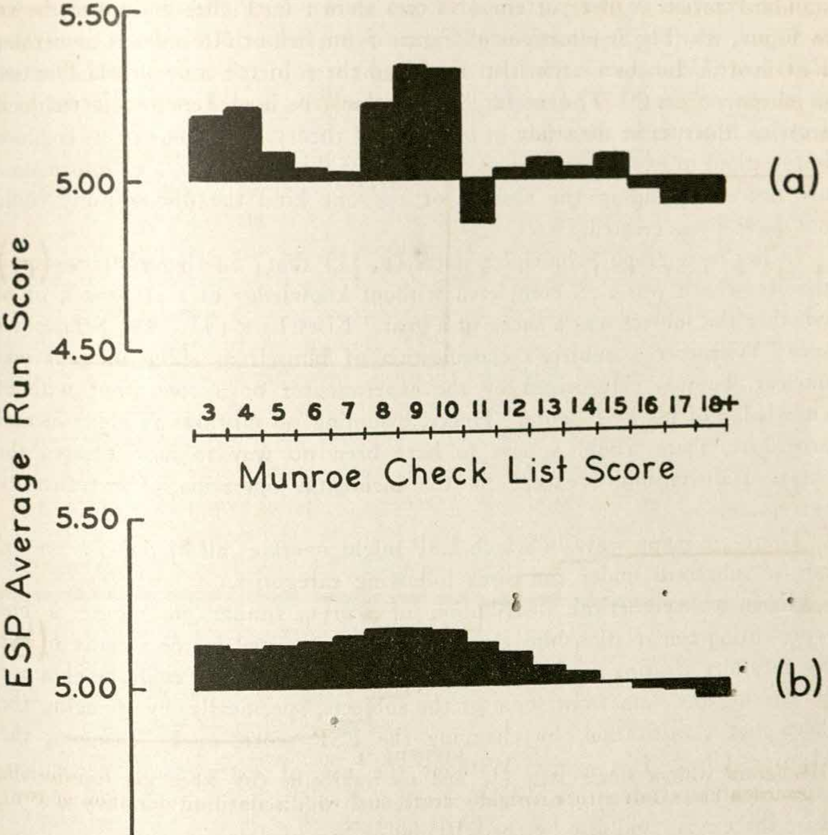


FIGURE 4
Observed average ESP scores for sheep (a) and their linear transform (b) when the transforming function has x-axis errors normally distributed with a standard deviation of 2. (Data from Table 10 of Schmeidler's monograph, grouping extreme scores at 3 and 18 and assuming every sheep had 215.5 ESP trials)

Check List score and ESP performance, and hence that in this sense the discontinuity must be an artifact.

Before examining Schmeidler's second possible explanation of the discontinuity, it may be well to consider in the broadest terms what this discontinuity implies and how it could have been created.[3]

The creation of the pattern of scores shown in Figure 1 implies the association, whether by sensory or extrasensory means, of information concerning ESP score, Munroe Check List score, and the subject's attitude classification as sheep or goat. (The word "information" is used here in its technical sense as it occurs in the study of information theory, with none of its common implications of explicit awareness, etc.) All of these three kinds of information are necessary, and in the absence of any one kind the discontinuity could not have been created.

In her monograph Schmeidler states (p. 11) that "Scoring and revision of the Rorschach was . . . completed without knowledge of ESP scores, or of whether the subject was a sheep or a goat." Elsewhere (4, p. 48) Schmeidler says "Whenever a subject's classification of himself as sheep or goat was unclear, he was categorized by the experimenter or a consultant without knowledge of his ESP score." Hence, assuming no mistakes in experimental procedure, there would appear to have been no way to have created the observed discontinuity except by the incidental operation of extrasensory perception.

There are many ways in which ESP might operate; all of these, however, can be subsumed under the three following categories.

Given a hypothetical distribution of scores, similar to Figure 1 but representing the relationships that would have resulted in the absence of the discontinuity-creating mechanism, the observed discontinuity could be obtained by altering the data from some of the subjects, specifically, by changing the sheep-goat classification, by changing the ESP score, or by changing the Munroe Check List score. Without the use of extrasensory perception, experimental peradventure might create such a discontinuity in any one of these three ways and also by the selection or loss of data.

As set forth elsewhere (4 p. 46 ff.) many precautions were taken against the loss or selection of data and against the accidental changing of the data once they were recorded. Moreover, a substantial sample of the original data records were examined (4 p. 123 ff.) by RAM. He verified the sheep-goat classifications by reading the questionnaires filled in by the subjects and he rescored the original ESP responses against the target lists. No errors of any consequence were found. It should be mentioned also that Figure 2 above suggests specifically that the ESP scores at the discontinuity deviate too far from chance to have been created by an occasional change of a miss or a hit.

---

[3] Schmeidler proposes and dismisses a third explanation (2 p. 30) of her discontinuity, namely that she assigned Munroe checks according to whether from personal contact she unconsciously recognized the subjects as types who were likely to be successful at the ESP test. From the point of view of the present discussion this is similar in principle to her second hypothesis.

Taken together, the evidence points most strongly to the possibility that the discontinuity was created by a shift of the Munroe Check List scores occurring somehow at the time of the scoring. A considerable looseness in the scoring of the Rorschach protocols (2 p. 10 ff.) points in the same direction but leaves unclear the details of the anomaly-creating mechanism.

Schmeidler's second possible explanation conforms to the foregoing considerations. Schmeidler states (2 p. 28 ff.) that the Rorschach "records were scored with more care when they were in the neighborhood of the previously set dividing line between 10 and 11 . . ." She goes on: "Did this additional concern result in some qualitative difference in my scoring of these borderline records? . . . I was curious enough about it to go to my files . . . Having piled up a substantial number of protocols originally scored with either 10 or 11 Check List entries, I tried to rescore them as if they were new. . . . Before a dozen records were rescored it was clear that all contained some doubtful category where more than a mechanical judgment was required. . . . The range of scores which it seemed to me might perhaps be defensible was two to five, in these first records, with a mode at three. [I.e., the most likely uncertainty was three score points *vide supra*] . . . . My next step was to sort the records into those which had been scored at 10 and those which had been scored at 11, and read over a number from each batch to see if any clear pattern emerged. My impression here was that the records which showed lively, impulsive vitality—sometimes misdirected, often associated with anxiety, aggression or guilt—had been scored at 10; and that the records indicating compulsive constriction and constraint had been scored at 11. This might be the key to the problem of why the pooled scores of 10 contrasted so markedly with the pooled scores of 11 . . . near the 10-11 boundary line my decisions may have been guided by the theory that more liveliness implied more flexibility and therefore better potential adjustment than constraint; and entries may have been assigned or withheld . . . according to this theory. If—as many previous studies have indicated—liveliness is associated with better ESP scoring than is constraint, the problem of why ESP scores at the 10 level were so much better than those at 11 has at least a possible answer."

The matter is not elaborated elsewhere in the monograph, but on page 38 it is rephrased as follows: "Re-examination of the scoring indicates that for many records which were immediately above or below the pre-set dichotomizing line, and for which there were some doubtful entries, the doubt was usually resolved in favor of good adjustment scores for the records which gave an over-all impression of liveliness, and in favor of poor adjustment for the records which gave an over-all impression of constraint." Similar statements appear on page 60 and in the Summary of the monograph (p. 66).

In other words, Schmeidler suggests (*a*) that the Rorschach records showed "liveliness" vs. "constraint," (*b*) that she used this criterion only near the 10-11 line, and (*c*) that this criterion is related to ESP performance and thus created the observed discontinuity.

This hypothesis, if true, is the most important finding of the entire research—judging from the sensitive way in which it distinguishes good and poor ESP subjects so as to create a huge statistical anomaly. It is unfortunate that after making her discovery Schmeidler did not submit all of her 10 and 11 score Rorschach records to independent judges to discover whether they could recognize the liveliness-constraint dichotomy which she found.

Instead, Schmeidler's procedure was to regard the pre-planned analysis utilizing the 10-11 dichotomy as "interim" (*ibid.* Chapter V) and to carry out a new "formal" analysis (Chapter VI) in which the Check List scores were tri-chotomized into groups 2-7, 8-12, and 13+. Schmeidler says (p. 39) that by the corresponding three-level analysis of variance " . . . the systematic variation [i.e., discontinuity] I suspect to have been present is presumably nullified . . ."

Aside from those proffered by Schmeidler, there are numerous possible mechanisms by which the observed discontinuity might be explained. One of these is that while scoring the Rorschach protocols Schmeidler, using ESP, was unconsciously aware of the sheep-goat classification as well as the (present or future) ESP scores of her subjects, and in marginal cases unknowingly assigned Munroe checks so as to "prove" her experimental hypothesis.

Schmeidler rejects this hypothesis (3) although no experimental controls were attempted against it. If one chooses to consider this kind of hypothesis, one must, *ipso facto*, be sorely perturbed by the difficulties facing the scientist in the field of parapsychology.

SUMMARY

The author examines critically a table from Gertrude R. Schmeidler's recent Rorschach monograph, confirming the occurrence therein of a non-

chance discontinuity in the relationship between average ESP score and the Check List score obtained by Munroe's Rorschach Inspection Technique. The author rejects as mathematically impossible the first of two explanations of this discontinuity offered by Schmeidler and believes that her second explanation deserves more serious attention than she has given it. He states that there are other conceivable explanations and cites the possibility that Schmeidler may herself have inadvertently used ESP in scoring the Rorschach protocols.

REFERENCES

1. McCONNELL, R. A. Scaled Measurement in Psi Research. *J. Parapsychol.*, 1958, **22**, 237-245.
2. SCHMEIDLER, G. R. ESP in Relation to Rorschach Test Evaluation. New York: Parapsychology Foundation, 1960. Pp. 89.
3. ———. (personal communication).
4. SCHMEIDLER, G. R., & McCONNELL, R. A. ESP and Personality Patterns. New Haven: Yale Univ. Press, 1958. Pp. 136.

Department of Biophysics
University of Pittsburgh
Pittsburgh 13, Pennsylvania

Published as a separate and in *The Journal of Psychology*, 1961, 52, 99-104.

ORGANIZATION OF RELEVANT AND IRRELEVANT WORDS IN THE SOLUTION OF VERBAL PROBLEMS*[1]

Department of Psychology, Michigan State University

D. M. JOHNSON AND E. R. HALL

A. INTRODUCTION

This paper assumes that the solution of many problems begins with some kind of organization of the data presented and reports two experiments on the difficulty of such organization. When the solution of a problem is specified by various numbers of unorganized items, either words or digits, there is a clear relation between the number of items and the time required for organizing them (3). The experiments to be reported here attempt to describe this organizing process in more detail by including irrelevant words along with the relevant.

The influence of irrelevant data is not usually made explicit in research on problem solving but this variable has been considered in some recent work on learning. In an experiment on acquisition of figure concepts Archer, Bourne, and Brown (1) varied irrelevant information from one to five bits, holding relevant information constant at two bits, and found a large increase in errors and in time to criterion. In Restle's (4) theory and experiment on discrimination learning both relevant and irrelevant cues are considered and the proportion of relevant cues is assigned a central role in learning and transfer.

To clarify the role of irrelevant data in the solution of problems presented only once we need a variety of short problems in which the number of relevant and irrelevant items can be varied independently over a wide range. Verbal concept problems of the type illustrated below are particularly suitable for this purpose because the organizing operation can be separated from the rest of the problem-solving enterprise. The instructions are to observe what the four relevant words in the left column have in common, ignoring the two irrelevant words, and find another word of the same class in the column at the right.

---

* Received in the Editorial Office on February 16, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Supported by a grant from the National Science Foundation.

bark
salt
pepper
cement
paprika
thyme

1. birch
2. number
3. wood
4. pot
5. cloves

It is assumed that *S* studies the left column and somehow organizes the relevant words, eliminating the irrelevant words in the process, then turns to the right column and selects a word that fits this cognitive organization. The first operation could be called induction and the second deduction or, in more general terms, the first could be called preparatory to the second. Whatever names are used, it is possible to time the two operations separately by serial exposure (3). A 5 × 8 card, on which a problem is printed as above, is inserted in a serial-exposure box by which the two columns are illuminated separately. The left column is illuminated first, then, when *S* is ready, he turns a switch that shifts the light to the right column, selects a solution, and presses one of five numbered buttons. Thus for each problem a record is obtained of time spent on the left column, called preparation time, time spent on selection of a solution, called solution time, and the solution selected. The present experiment is directed primarily toward the time spent on the first or preparatory operation, and one of the advantages claimed for the serial-exposure method is that it permits timing this operation separately.

B. Experiment I. Relevant and Irrelevant Words

Problems of the type illustrated above were constructed of words assumed to be familiar to college students. Previous research with verbal analogies exposed serially has shown that it is possible, by using words of different levels of difficulty, to increase time spent on either the inductive phase or the deductive phase (2). In the construction of problems for the present experiment the intent was to locate the difficulty in the preparatory organization (induction) and make the selection of a solution (deduction) relatively easy.

The number of relevant words, n, was varied from two to eight, and the number of irrelevant words, n′, from zero to six, and one problem of each combination of n and n′ was constructed, making 49 problems in all. Thus the total number of words, n + n′, varied from two to 14. Relevant and irrelevant words were mixed irregularly in the left column, and the position of the correct word in the right column was randomized. On another set of cards the same problems were printed in reverse order, and each set was used for half the *Ss*. The cards were shuffled between *Ss* but were always preceded by an introductory period during which six practice problems were done and the instructions were clarified if necessary. Speed was not stressed but most of the *Ss* knew they were being timed. Each *S* did all 49 problems in two periods of 20 to 30 minutes each, one or two days apart. The 20 *Ss* were recruited from an elementary psychology class.

The number of errors was small, about two per *S*, and they showed no apparent relation to n or n' or preparation time. The important data are the 49 median preparation times shown in Table 1, which have an over-all mean of 5.9 and a standard deviation of 2.1 seconds. The effect of the irrelevant words shows clearly in the table, but the effect of the relevant words is obscure. The product-moment correlation with preparation time is .86 for n' and —.01 for n.

Closer study of Table 1 and consideration of *S*'s task leads to a more rational description of the effects of these two variables and of their interaction. It is reasonable to suppose that some time is spent reading each word, and a linear function fitted to the bottom line (relevant words only) of Table 1 suggests about 0.2 second per word. Extrapolation of this function to include irrelevant words as well would represent the time spent in simple reading of each word of both classes. The correlation between preparation time and $n + n'$ is .60.

TABLE 1
MEDIAN PREPARATION TIMES IN SECONDS FOR VERBAL PROBLEMS OF 2–8 RELEVANT WORDS AND 0–6 IRRELEVANT WORDS

Irrelevant words	Relevant words						
	2	3	4	5	6	7	8
6	9.8	6.7	8.4	7.7	8.3	6.7	7.3
5	9.3	7.5	7.4	9.3	9.6	7.2	5.7
4	7.4	7.4	5.3	5.8	8.3	6.5	9.0
3	4.6	6.9	7.0	7.7	5.9	5.5	5.6
2	4.1	4.4	6.3	7.1	5.0	5.3	4.9
1	3.3	4.8	3.5	3.6	3.5	4.0	4.2
0	2.4	2.2	2.7	2.8	3.3	3.5	3.2

The special effect of the irrelevant words is probably proportional to the total number of words. Six irrelevant words are more disturbing when accompanied by only two relevant words than when accompanied by eight relevant words. The critical quantity would appear to be $n'/n + n'$, which correlates .79 with preparation time, with no evidence of nonlinearity. These effects may be combined in an equation, written in seconds, using P for median preparation time. $P = .3(n + n') + 7.8(n'/n + n') + .7$. This equation reproduces the obtained medians of Table 1 with a correlation of .87 and a standard error of estimate of 1.1 seconds.

More precise fit should not be expected because there is only one problem of each combination of n and n′ and the difficulty of the single words is uncontrolled. If we accept this equation as accounting for most of the time spent in organizing these word lists preparatory to selecting a solution, several statements of theoretical significance can be made. The three terms of the equation indicate that the preparatory time can be divided into reading time which is the same for all words, a special effect of the irrelevant words, and a constant which presumably measures the time spent in getting started on each problem and turning the switch to the solution exposure. The special effect of the irrelevant words, which is proportional to the total number of words, accounts for much more of the variance of preparation time than simple reading time. (The correlation between $n + n'$ and $n'/n + n'$ in this set of problems is .32.) The high correlation between n′ and preparation time is due to the fact that the irrelevant words contribute to reading time as well as to the special irrelevant effect. The negligible correlation between n and preparation time is due to the fact that the contribution of the relevant words to reading time is balanced by their reciprocal influence in reducing the effect of the irrelevant words.

Solution time, as defined above, was also recorded, and the 49 medians, corresponding to the 49 median preparation times were computed. The over-all mean of these medians was 2.7, with a standard deviation of .3 second. These 49 medians were correlated with other variables of interest, with the results shown in Table 2. The correlation of only .35 with preparation time justifies timing each process separately, and the other correlations demonstrate that solution time is not determined by the same factors that determine preparation time. The correlations in the bottom line of Table

TABLE 2

CORRELATION COEFFICIENTS SHOWING THE CONTRIBUTION OF NUMBER OF RELEVANT WORDS, N, AND NUMBER OF IRRELEVANT WORDS, N′, TO TIME SPENT ON VERBAL REASONING PROBLEMS

	Preparation time	Solution time	n	n′	n + n′	$\frac{n'}{n + n'}$
Preparation time		.35	—.01	.86	.60	.79
Solution time	.35		—.21	.18	—.02	.29
Total time	.99	.46	—.05	.84	.56	.78

2 show that preparation time was the chief ingredient of total time or, in other words, that the attempt to construct problems in which the preparatory organization was the principal difficulty was successful.

C. EXPERIMENT II. RELATED AND UNRELATED IRRELEVANT WORDS

The irrelevant words of Experiment I were unrelated to each other as well as to the relevant words, but it is a plausible hypothesis that irrelevant words which are related to each other in such a way that they can be easily organized into a competing conceptual pattern will be more disturbing. To test this hypothesis some of the problems used in Experiment I were modified by replacing the irrelevant words with others that could easily be related to each other. For example, the illustrative preparatory list printed above was changed to stone, salt, pepper, cement, paprika, thyme. It was assumed that the related irrelevant words stone and cement would interfere more with the organization of the relevant words than the unrelated irrelevant words bark and cement. Problems of four to eight relevant words and two to three irrelevant words, two of each combination, making 20 in all, were constructed in two forms, one with the irrelevant words unrelated and the other identical except that the irrelevant words could easily be organized in a group of two or three words. The 20 problems of one form were given to 10 *S*s and the parallel 20 problems in the other form were given to 10 other *S*s. The details of procedure were approximately as in Experiment I and were the same for both groups.

Mean errors for the related irrelevant words were 1.4 per *S*, for the unrelated irrelevant words 1.2 per *S*. The data on median preparation time are shown in Table 3, and it is clear that there is also no significant difference in time. Either the hypothesis is wrong or we have not tested it properly.

TABLE 3
MEDIAN PREPARATION TIMES IN SECONDS FOR VERBAL PROBLEMS OF 4–8 RELEVANT WORDS AND 2–3 RELATED AND UNRELATED IRRELEVANT WORDS

	Relevant Words				
	4	5	6	7	8
Related Irrelevant Words					
3	8.0	6.6	7.2	7.2	7.3
2	6.7	5.3	6.8	5.6	6.1
Unrelated Irrelevant Words					
3	7.3	8.5	6.6	7.2	7.1
2	5.4	6.3	6.2	5.3	7.6

The data of Table 3 can be used as a limited cross validation of the equation formulated from the data of Experiment I. Since there were no differences of any significance between the top and bottom halves of Table 3, the two halves were combined and the 10 means compared with the corresponding values estimated by the equation printed above. This comparison yields a standard error of estimate of 1.2 seconds, about the same as for the original data. Since the *S*s and most of the problems were different, this finding adds some support to the theoretical statements above.

D. SUMMARY

It was assumed that the solution of many problems begins with some sort of organization of the data, preparatory to actual selection of a solution, hence Experiment I investigated the effects of relevant and irrelevant items on the difficulty of such preparatory organization. Verbal concept problems were constructed so the preparatory material and the alternative solution possibilities could be exposed serially by *S* and the time spent on each operation could be recorded separately. There were 49 problems of two to eight relevant words and zero to six irrelevant words to be organized. Analysis of the median preparation times for 20 *S*s indicated that some time was spent in reading words of both classes, $n + n'$, but that most of the variance was due to the proportion of irrelevant words to all words, $n'/n + n'$. An equation that combines these effects predicted the 49 obtained medians with a correlation of .87 and a standard error of estimate of 1.1 seconds.

Experiment II began with the hypothesis that irrelevant words that can be easily related to each other interfere more with the organization of the relevant words than unrelated irrelevant words. This hypothesis was not confirmed, but the data obtained permitted application of the equation formulated from the data of Experiment I, and the resulting standard error of estimate was 1.2 seconds. This limited cross validation adds further support to the above evaluation of relevant and irrelevant data.

REFERENCES

1. ARCHER, E. J., BOURNE, L. E., & BROWN, R. G. Concept identification as a function of irrelevant information and instructions. *J. Exp. Psychol.,* 1955, **49**, 153-164.
2. JOHNSON, D. M. Serial analysis of thinking. *Annals N. Y. Acad. Sci.,* 1961, **91**, 66-75.
3. JOHNSON, D. M., LINCOLN, R. E., & HALL, E. R. Amount of material and preparation time for solving verbal problems. *J. Psychol.,* 1961, **51**, 457-471.
4. RESTLE, F. A theory of discrimination learning. *Psychol. Rev.,* 1955, **62**, 11-19.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 105-114.

THE VARIATION OF CHOICE REACTION TIME AMONG GROUPS OF HOSPITALIZED PSYCHIATRIC PATIENTS*

Psychiatric Institute, Columbia University

DONALD J. DILLON[1,2]

A. INTRODUCTION

The study of psychomotor behavior characterized the beginning of psychology as a science in its own right. Since then, such types of motor activity as reaction time, tapping behavior, manual dexterity, coordination, etc., have occupied a great deal of the attention of psychological investigators. Even the current issue of *Psychological Abstracts* contains abstracts of some type of psychomotor investigation.

Reports of investigations into the psychomotor aspects of the mentally ill were, until recently, relatively rare. Such investigations as do exist, however, may be considered to fall into two broad categories: those dealing with the impaired performance of the mentally ill patients and those dealing with the influence of the variation of the neurologic structures and of the physiology on psychomotor performance.

The investigations of Huston and Shakow (2, 3, 4, 12, 13) are particularly noteworthy in connection with the impaired performance of the psychotic patient. These authors investigated reaction time, tapping speed, steadiness and work rates in schizophrenic and manic-depressive patients. These patients did display impaired performance. The impairment was attributed to the lowered motivational status of these patients.

Windle and Hamwi (14) administered a complex reaction time test to chronic and early psychotic patients. The early group averaged faster times than did the chronic group. Within the early group there was a positive

---

* Received in the Editorial Office on February 24, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was made possible by funds granted by the Carnegie Corporation of New York and by grants M-872 and MY-3616 from the National Institutes of Health, USPHS.

[2] The author is indebted to Miss Ruth Link and Miss Jane Courten for assistance in the statistical analysis, to Dr. Carney Landis for his critical reading of the manuscript, and to Dr. Joseph F. Kubis for his very helpful suggestions as to the appropriate statistical techniques.

relationship between score and clinical outcome; within the chronic group, the relationship was inverse.

King (5) reported an extensive investigation of reaction time (lift and disjunctive) along with other measures of motor performance as they relate to three groups of patients: chronic schizophrenes, pseudoneurotic schizophrenes, and neurotics. A group of normal subjects was also tested.

King found age and sex differences in all motor performance scores obtained from his normal group. He also found the "usual" practice or learning curves, i.e., performance improved to a stable level. The chronic schizophrene group did not show the pronounced age and sex differences but did show the normal-like learning curves. The absolute measures for the group, however, were considerably below those of the normal group. The subacute group (pseudoneurotic schizophrenes and neurotics) were faster than the chronic group but slower than the normal group. Most of the deviation of this subacute group stemmed from the pseudoneurotic schizophrenes.

Psychomotor performance has also been investigated during periods immediately following brain surgery (prefrontal lobotomy) and following electroconvulsive, insulin and drug therapies. The results of such investigations are somewhat inconclusive, or, in some cases, controversial.

King and Clausen (6) found no impairment of performance on tapping, response time or finger dexterity tests following frontal lobe thermocoagulation, transorbital lobotomy or topectomy enlargement. They did find impaired performance, however, after ligation of the cerebral superior veins. In a subsequent study (7), the same two authors found a highly significant impairment of performances on all three tests following orbital or superior ablation of the frontal cortex.

Landis and Clausen (8) found an increase (impairment) both in decision and ballistic reaction times following prefrontal lobotomy (OP) and a course of insulin coma therapy (ICT), but not after a course of electroconvulsive therapy (ECT). Landis, Dillon, and Leopold (9), on the other hand, found such an increase in both decision reaction time and choice reaction time (CRT) following a course of ECT. Landis, Dillon, Leopold, and Rutschmann (10) did not find a significant impairment of CRT performance during or following a course of ICT, but did find a positive relationship between level of performance and blood sugar level.

In addition to the effects of therapy, Landis and Clausen (8) also reported that both decision and ballistic reaction times varied with age (younger were faster), with sex (males were faster), with diagnosis (psychoneurotics were faster, schizophrenics slower and depressed slowest) and with outcome (those who later showed the most clinical improvement were slowest). These relationships were demonstrated for the pre-treatment scores.

Landis, Dillon, and Leopold (9) and Landis, Dillon, Leopold, and Rutschmann (10) found no significant relationship between outcome and CRT score. No attempt was made to relate CRT to age or diagnosis in either of these two investigations.

B. Purpose

The CRT scores have been shown to relate to age, sex, diagnosis, and (possibly) outcome and also to reflect the effects of psychiatric treatment. Because of these findings, the author decided to collate all the data obtained in five previous investigations (6, 8, 9, 10, 11), and, having obtained such a large sample of patients, reanalyze in an attempt to shed some light upon the possibility of employing CRT as a prognostic indicator, while taking into account the effects of age, sex, and diagnosis upon the CRT. The reanalysis was also attempted in order to determine if the various forms of psychiatric treatment differed in their effects upon psychomotor performance, and to compare the results of these investigations with those obtained by King (5).

C. Apparatus and Procedure

The various pieces of equipment employed to measure the CRT have been described in previous publications (6, 8, 9, 10). The apparatus employed to measure CRT as influenced by the various phenothiazine derivatives (11) has not been described previously but the equipment was the same as that employed in the investigation of the effects of ECT (9) with one major modification. The ballistic and decision reaction time recording circuits were combined in order to yield a single CRT score.

For this reanalysis, the decision and ballistic reaction times in the Landis and Clausen study (8) were summed to yield a single CRT score. The scores obtained in the drug study (11), in the Landis, Dillon, and Leopold study (9) and in the Landis, Dillon, Leopold, and Rutschmann study (10) were already in terms of CRT. In order that the King and Clausen scores (6) conform as closely as possible to these other scores, the visual response times from their report were selected.

The scores selected for this analysis were those obtained immediately prior to the beginning of a given treatment (Pre), those obtained as soon as possible after treatment was begun and while it was still in progress (Dur), and those obtained from two to four weeks after the last treatment (Post).

Because of differences in apparatus, direct inter-experimental comparisons were impossible. For this reason, Z transform equations were computed for each of the five distributions, and each score was then converted to its corresponding Z unit. In this way, the means and standard deviations for each experiment were equated to 50 ("centiseconds") and 10 ("centiseconds") respectively, and differences in apparatus did not constitute a determinant of the CRT score.

Since changes in performance were brought about by treatment and/or practice, Z transforms were based upon the pre-treatment scores in order that such differences not be lost in the transformation. Three Z scores were computed for each patient whenever the patient was tested in all three sessions. All patients were not tested three times because of such things as early discharge, confounding of treatment, equipment failure, etc.

The differences in scores were then computed in such a way that a positive algebraic sign (+) was given to an improvement in performance. The computations were in the directions of: Pre *minus* Dur, Dur *minus* Post, and Pre *minus* Post.

The Z scores and changes in score were then classified as follows:

(*a*) *Treatment:* Phenothiazine derivatives (DRUG); psychotherapy (PSA); insulin coma therapy (ICT); electroconvulsive therapy (ECT); prefrontal lobotomy (OP). The phenothiazine derivatives were thorazine, sparine, vesprin and compazine.

(*b*) *Age:* 16-25; 26-45; 46 and older.

(*c*) *Sex.*

(*d*)*Diagnosis:* Dementia praecox paranoid (P)—the largest schizophrenic sub-group; dementia praecox (DP), exclusive of paranoid and depressed; depressed (D)—included all patients whose major symptom at the time of admission was one of depression.

(*e*) *Outcome:* 1. Much improved—patients exhibiting marked improvement. Most rated as such were later admitted to convalescent care status or discharged outright; 2. Improved—patients exhibiting some or noticeable improvement. A very few were later placed on convalescent care status; 3. Unimproved—patients showing no noticeable improvement. Most of these patients remained in the hospital.

A multivariate analysis of variance was impossible because of the lack of entries in many cells. Accordingly, a series of one-way analyses of variance were performed—testing one of the five variables at a time and one session at a time. The pre-during, during-post, and pre-post changes were analyzed in the same way.

D. RESULTS

The distributions of patients according to age, sex, diagnosis, outcome and treatment, and the first and second order interrelationships of these variables have been described in a prior publication (1). Although the number of patients herein described differed (204 vs. 250), the relationships and the extent of these relationships were the same as those described earlier. They will not be repeated here.

The intersession correlation coefficients (Pearson) were as follows: Pre:-Dur=+0.63; Dur:Post=+0.52; Pre:Post=+0.55. These coefficients (of reliability), while statistically significant ($p<.001$) are low. Since some form of treatment intervened, this is not unexpected.

The results of the analysis of the effects of main variables upon the CRT scores were as follows: (See Table 1).

(1) *Treatment.* CRT scores varied significantly according to treatment in the Dur session ($p<.01$) and for the Post session ($p<.02$). Those patients treated with psychotherapy had the fastest reaction times and those who were operated upon or who received ECT had the slowest times while treatment was still in progress. After treatment was terminated the groups of PSA patients and ECT patients were fastest and the groups of DRUG and OP patients were slowest.

(2) *Age.* CRT scores varied directly with age (younger were faster) for the Pre ($p<.02$) and Dur ($p<.01$) sessions. The scores did not differ with age after treatment was terminated. The correlation coefficient for age and Pre-treatment CRT was +0.22 ($p<.04$) while those for the During session was +0.26 ($p<.01$). The Post-treatment correlation was +0.03.

(3) *Sex.* Before treatment, males were significantly faster than females ($p < .05$). While treatment was in progress this sex difference was somewhat greater ($p < .01$). The sex differences disappeared after treatment was stopped.

(4) *Diagnosis.* The group of depressed patients showed slower CRT scores than the other two diagnostic groups before treatment was begun ($p < .01$). This difference in CRT was maintained but at a lower level of significance ($p < .05$) while treatment was in progress, and disappeared after the termination of treatment.

(5) *Outcome.* Differences in CRT as a function of outcome rating were not significant, except for the Post treatment session ($p < .05$). In this case, patients with an outcome rating of 1 were faster than those whose rating was 2 ($p <.01$).

The changes in CRT scores from the Pre- to Dur- to Post-treatment session reflect the effects of treatment (Pre-Dur) and the recovery from these effects (Dur-Post). These difference scores were subjected to one-way analyses of variance with these results:

1. *Treatment.* This variable exerted a significant effect in the Pre-Dur (p < .05) and Dur-Post (p < .01) comparisons. ECT and OP were followed by the greatest loss (Pre-Dur) and also by the greatest gain (Dur-Post). Treatment was not significant in the Pre-Post comparison.

TABLE 1

The numbers of patients and means and standard deviations of CRT according to treatment, age, sex, diagnosis and outcome, for each session. Values are Z score units and may be considered in terms of centiseconds. Note that scale is inverse: the higher the value, the poorer the performance.

1.	Treatment		DRUG	PSA	ICT	ECT	OP		
	Pre	N	17	23	58	76	30		
		$\bar{x}$	49.06	47.33	49.07	50.99	51.51		
		SD	8.80	8.19	10.24	10.13	10.29		
	Dur	N	16	23	56	76	15		
		$\bar{x}$	50.11	44.68	50.48	54.74	59.00		
		SD	11.77	7.93	10.20	13.16	14.75		
	Post	N	8	18	35	67	13		
		$\bar{x}$	53.64	44.88	47.43	45.86	51.83		
		SD	14.69	8.76	7.45	6.52	7.29		
2.	Age		16-25	26-45	46+	3. Sex		M	F
	Pre	N	65	100	39			108	96
		$\bar{x}$	47.78	49.86	53.79			48.34	51.75
		SD	9.67	8.90	11.37			9.94	9.57
	Dur	N	60	94	32			99	87
		$\bar{x}$	48.84	52.41	57.63			50.82	53.68
		SD	11.43	11.55	16.27			13.16	12.04
	Post	N	43	71	27			82	59
		$\bar{x}$	46.72	47.32	47.21			46.66	47.58
		SD	8.63	8.13	6.60			7.55	9.45
4.	Diagnosis		P	DP	D	5. Outcome	1	2	3
	Pre	N	59	109	36		75	78	51
		$\bar{x}$	50.49	48.17	54.44		50.27	50.62	48.44
		SD	10.32	8.78	10.85		9.46	10.46	9.52
	Dur	N	53	97	36		72	71	43
		$\bar{x}$	51.75	50.37	57.58		52.07	52.65	51.48
		SD	11.43	11.53	15.73		10.71	13.24	14.82
	Post	N	41	70	30		60	50	31
		$\bar{x}$	46.33	47.30	47.75		45.30	49.46	46.86
		SD	9.41	7.23	7.82		6.26	9.27	7.92

2. *Age.* Age proved to be a significant factor in the Dur-Post (p < .01) and Pre-Post (p < .05) comparisons. In both comparisons, the oldest patients registered the greatest gains in score.

3. *Sex.* No significant effects of sex were found.

4. *Diagnosis.* Significant differences in terms of change in CRT score were found in the Dur-Post ($p < .01$) and Pre-Post ($p < .05$) comparisons. It was the depressed group of patients who gained the most.

5. *Outcome.* No significant differences were found.

E. DISCUSSION

Several things must be pointed out before these results can be discussed. The first is a matter of terminology. The CRT as used here is in line with the choice reaction time as discussed by Woodworth (15) in that there was one response (move and touch a plate) called for by either of two lights. This also corresponds to King's (5) disjunctive reaction time (jump).

Secondly, the clinical groups are necessarily vague. The investigations were conducted in several State institutions and in institutions in two different states and were spread over a 10 year period. The author had to work from case histories and the diagnostic descriptions contained therein. Consequently, the groupings here were obtained by placing those patients presenting somewhat similar clinical pictures into three (as it turned out) groups. It was virtually impossible to delineate these groupings more sharply.

This reanalysis has demonstrated, quite clearly, the relationship of CRT to psychiatric treatment, age, sex, and diagnosis. The during scores, the Pre-Dur and Dur-Post comparisons all revealed the significant effects of the treatment variable in that ECT and OP were followed by a large decrement in performance. Age, sex, and diagnostic differences were demonstrated by the Pre- and During-treatment analyses and by the Dur-Post and Pre-Post comparisons. The older patients and the depressed displayed the slowest CRT scores, both before treatment was begun and while it was in progress. These same groups of patients also showed a greater gain after termination of treatment (Dur-Post) and also a greater "overall" gain (Pre-Post). The male patients were faster than the females before and during treatment, but there were no significant sex differences in the inter-session comparisons. Age, sex, and diagnostic differences disappeared after termination.

There was no effect of treatment grouping upon the pre-treatment CRT scores, i.e., no pre-selection of CRT scores with respect to treatment was in evidence. Consequently, the significant differences in CRT scores for the during session with respect to treatment was a treatment effect *per se;* they did not occur because of selection. These treatment differences (Dur) coupled with the Pre-Dur and Dur-Post differences demonstrated impairment in performance as a result of ECT and OP. After termination of ECT or suitable time lapse following surgery, the impairment was no longer in evidence.

CRT increased with age both before and during treatment. After termination, the relationship disappeared. This disappearance, together with the reduction in the magnitude of the SD in the Post session, would point toward a regression effect. However, the significance of the age effect upon the Dur-Post and Pre-Post changes (older patients gained more than did younger) demonstrated that something more than the usual regression effect occurred. Apparently, in recovering from the effects of treatment, all patients registered gains in performance, but the oldest group of patients showed the greatest degree of recovery—significantly greater than the two younger groups. In other words, as the older patients recovered (in terms of CRT) from the effects of treatment, they also tended to overcome the initial performance "handicap" which was associated with their greater age.

Before and during the course of treatment, male patients registered faster scores than did the female patients. The differences disappeared after termination. Such sex differences have been found previously (King—5). The disappearance was a manifestation of the regression effect since the intersession comparisons revealed no significant sex effects.

Diagnostic differences were manifested before and during treatment and were attributed to an impairment which accompanied depression. Since the diagnostic differences disappeared upon termination, and the depressed patients registered the greatest gains (Dur-Post), it may be said that the depressed patients behaved like the older patients, i.e., in recovering from the effects of treatment, the depressed patients also compensated for their initially poorer performance.

The significance of the outcome effect (barely .05) is actually unimportant, since the significance can be attributed to a selection factor. Only 67 per cent of the "2" group gave post-treatment scores. The average initial CRT for the group which "dropped out" was faster than that of the group which was tested in the Post session.

Finally, since King (5) has presented an extensive survey of the psychomotor performances of psychiatric patients, a comparison of the results of this analysis with those of King's survey is in order. Briefly, King found a clear manifestation of an age factor and the minimal presence of a sex factor. The performances of the subacute group (particularly pseudoneurotic schizophrenics) were retarded in comparison with the normal group; the performances of the chronic group (schizophrenics) were retarded when compared with those of the subacute group. The practice curves were similar for all three groups. Within the chronic group there was a close correspondence between psychomotor performance and clinical status; there seemed to be no relationship between motor performance and either age or sex within this group.

The performances of the patients analyzed in this current investigation are not strictly comparable with King's chronic group, but since enough similarities do exist, some comparisons may be drawn. Where King obtained no age or sex differences in his chronic group, such differences were found in our analysis—before and during treatment. King found a relationship between performance and clinical status, i.e., poorer performance was associated with a more disturbed symptomatic picture. In our case, poorer performance was associated with the presence of symptoms of depression (possibly the more disturbed patients). Since King had no post-treatment data (this was not his purpose), no comparison with our post-treatment scores may be made.

The data analysis in conjunction with the comparison with King's conclusions points to two conclusions with respect to the age and/or diagnostic differences revealed in this report. The similarity of our diagnostic differences to King's differences, attributable to clinical status, would argue for the existence of diagnostic differences *per se*. In opposition to this is the high positive relationship between age and CRT scores and the great degree of relationship between diagnosis and age (depressed are older). Unfortunately, a proper analysis of the interaction effects of age and diagnosis upon the CRT was impossible, so this problem cannot be resolved. Obviously, only further research and proper experimental design can give us the answer.

F. SUMMARY

Measurements of the choice reaction time had been obtained previously from five groups of hospitalized psychiatric patients. The data have been analyzed further in order to determine the effects of age, sex, diagnosis, type of therapy, and final disposition of the patient (outcome). The changes in threshold before, during, and after a course of prescribed treatment were also analyzed for these effects. The analysis pointed to these conclusions:

1. CRT measures obtained before a course of psychiatric treatment were significantly related to age, sex and diagnosis.
2. Those CRT measures obtained during a course of psychiatric treatment were influenced significantly by treatment, age, sex, and diagnosis.

3. Only the treatment variable proved to exert a significant effect upon CRT scores after termination.

4. Recovery from the effects of treatment (During-Post comparison of CRT scores) was also influenced significantly by age, diagnosis, and treatment.

5. The effects of treatment were significant when the pre- and during-treatment scores were compared.

REFERENCES

1. DILLON, D. J. The variation of flicker-fusion thresholds among groups of hospitalized psychiatric patients. *J. of Psychol.*, 1961, **51**, 351-360.
2. HUSTON, P. E. Eye-Land coordination in schizophrenic patients and normals as measured by the pursuit meter. *Psychol. Bull.*, 1932, **29**, 662.
3. HUSTON, P. E., & SHAKOW, D. Studies of motor function in schizophrenia: III. Steadiness. *J. Gen. Psychol.*, 1946, **34**, 119-126.
4. HUSTON, P. E., SHAKOW, D., & RIGGS, L. A. Studies of motor functions in schizophrenia: II. Reaction time. *J. Gen. Psychol.*, 1937, **16**, 39-82.
5. KING, H. E. Psychomotor Aspects of Mental Disease. Cambridge: Harvard Univ. Press, 1954.
6. KING, H. E., & CLAUSEN, J. Chapter 13 in F. A. Mettler (*Ed.*) *Psychosurgical Problems*. New York: Blakiston, 1952. (Pp. 254-274.)
7. ———. Chapter 9 in N. D. C. Lewis, C. Landis, & H. E. King (*Eds.*) *Studies in Topectomy*. New York: Grune & Stratton, 1956. (Pp. 126-149.)
8. LANDIS, C., & CLAUSEN, J. Changes in sensory motor perfomances induced by active psychiatric treatment. *J. of Psychol.*, 1955, **40**, 275-305.
9. LANDIS, C., DILLON, D., & LEOPOLD, J. Changes in flicker-fusion threshold and in choice reaction time induced by electroconvulsive therapy. *J. of Psychol.*, 1956, **41**, 61-80.
10. LANDIS, C., DILLON, D., LEOPOLD, J., & RUTSCHMANN, J. Changes in the level of blood sugar and sensory and motor performances brought about by insulin coma therapy. *J. of Psychol.*, 1958, **45**, 275-285.
11. LANDIS, C., DILLON, D., RUTSCHMANN, J., & LINK, R. Changes in sensory and motor performance brought about by phenothiazine treatment. (Unpublished data.)
12. SHAKOW, D. A study of certain aspects of motor coordination in schizophrenia with the prod meter. *Psychol. Bull.*, 1932, **29**, 661.
13. SHAKOW, D., & HUSTON, P. E. Studies of motor function in schizophrenia: I. Speed of tapping. *J. Gen. Psychol.*, 1936, **15**, 63-108.
14. WINDLE, C., & HAMWI, V. An exploratory study of the prognostic value of the complex reaction time as a test in early and chronic psychotics. *J. Clin. Psychol.*, 1953, **9**, 156-161.
15. WOODWORTH, R. S. Experimental Psychology. New York: Holt, 1938.

Psychiatric Institute
Columbia University
722 West 168 Street
New York, New York

Published as a separate and in *The Journal of Psychology*. 1961, **52**, 115-124.

VERBAL CONDITIONING: DISCRIMINATION AND AWARENESS*

Department of Psychology, Purdue University

FREDERICK H. KANFER AND JOHN F. MCBREARTY[1]

A. DISCRIMINATION

In recent studies on the effectiveness of reinforcement on verbal behavior almost no attention has been devoted to the nature of the discriminative stimulus (S^D) in whose presence reinforcement is said to result in "verbal conditioning." Furthermore, the usual analyses in these studies have assumed that *S* learns a new response class, as evidenced by increased probability of occurrence. Recent criticisms from various quarters (Dulany, 1960; Matarazzo, Saslow, and Pareis, 1960) have suggested that verbal conditioning may be a much more complex process. Research in our own laboratories has suggested that several processes may simultaneously go on during conditioning. Among these, discrimination learning, i.e., learning to identify and differentiate the critical stimulus dimensions, may become one crucial determinant of performance.

Especially with the construct-a-sentence-method first used by Taffel (1958), repeated negative findings (Krasner, 1958) indicate that the complexity of this paradigm makes a simple analogy to animal operant conditioning difficult to maintain. In the Taffel technique, *S* is usually shown several words of which he selects one. If it is a member of a predetermined critical class, e.g.: a pronoun referring to oneself such as I-We, a hostile or active verb, or an affective adverb, it is followed by emission of a verbal reinforcing stimulus by *E*. In this paradigm the range of critical responses is limited to the words shown on each trial. An *S* will show an increase in response frequency when he discriminates among the presented stimulus classes, since subsequent verbal responses (sentences) are always reinforced when one class of words (S^D) is used and never reinforced when another class (S^Δ) is used. Therefore, the exercise of considerable stimulus control in the Taffel procedure leads to

---

* Received in the Editorial Office on March 1, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was supported in part by Research Grant M-2027 (C3) from the National Institute of Mental Health, United States Public Health Service.

the expectation that the difficulty of the stimulus discrimination and the particular stimuli selected would affect learning to a greater extent than in free-operant conditioning where relatively little stimulus control is possible.

Further support for the suggestion that verbal conditioning is not a simple case of operant conditioning comes from various studies which indicate that awareness may modify the ease with which the designated response class can be increased. To date, there has been no demonstration however, that increased learning is a direct function of ability to verbalize the nature of the task rather than a manifestation of a simple correlation between performance and a later verbal evaluation of one's own performance.

Awareness has generally been used to refer to verbal response classes which specify the relationship between a critical response class and a reinforcing stimulus. Since the Taffel procedure restricts responses to the same semantic form as the S^D, studies using this procedure usually extend the term to cover also the specification of the S^D for emission of a reinforced response. Here awareness will be used to indicate correct verbalization of the response-reinforcement contingency.

The relationship between stimulus dimensions, verbalization of the reinforcement-response contingency, and learning can be postulated in two ways. First, "awareness" can be viewed as a mediating response occurring under specifiable stimulus conditions which directly affects learning. A second alternative is to consider that verbalization of the response-reinforcement contingency may be independently affected by *different* antecedent conditions than the acquisition rate of the response class, even though the experimental conditions may sometimes facilitate an interaction between these two events. This view implies that report of the correct response-reinforcement relationship may not always correlate with learning.

In the present study the latter view was taken. It was hypothesized that the presentations of S^Ds which vary in similarity (discriminability) from alternate choices (S^Δ) may separately affect *a*) the frequency of reported "awareness" in post-experimental interviews, and *b*) the rate of acquisition. Variation of S^D and S^Δ on a dimension of semantic similarity was expected to result in greater frequency of reported awareness under conditions of minimal stimulus similarity. Furthermore, it was hypothesized that maximum disparity between the discriminanda would also lead to the greatest increase in response probability during acquisition.

B. METHOD

The general procedure was similar to that developed by Taffel (1955) and used by Buss and Durkee (1958). The stimulus words were taken from lists prepared by Buss and used by Buss and Durkee (1958) and Binder, et al. (1957). The lists were constructed on the basis of ratings of the degree of hostile connotations. Three such lists were available, all matched in Thorndike-Lorge word count frequency. The word pairs were typed on 5 × 7 white cards. For each *S*, 20 word pairs were presented in random order constituting one block of trials. On each subsequent block order of presentation was again randomized. In order to facilitate administration a pre-arranged deck of 120 cards was used.

1. *Subjects*

Eighty-three male undergraduate *S*s were recruited from elementary psychology classes. They were randomly assigned to one of the two experimental groups. The data from two *S*s were discarded because they reported participation in a similar experiment at an earlier time. Data from one *S* were discarded because he consistently used words which appeared on one side of the card on 18 of 20 trials.

2. *Procedure*

Each *S* was asked to make up a sentence with one of the verbs appearing on the card. After 20 non-reinforced (pre-acquisition) trials, *E* reinforced use of the critical response class by saying "Good" at the end of the sentence. *S*s were divided into two groups. In group MH (mildly hostile) the word pairs on each trial consisted of one neutral word paired with a matched word from the *mildly hostile* list. Position of the words and their combinations were randomly varied. In groups IH (intensely hostile) the word pairs consisted of one neutral word paired with a word selected from the *intensely hostile* list. All pairs were matched on a Thorndike-Lorge frequency count.

Throughout the experiment *E* attempted to maintain a constant attitude and no verbal statements were made except those which served as reinforcing stimuli for critical responses. At the end of 120 trials *S* was asked to fill out a questionnaire with *E*, consisting of six questions. The questions attempted to ascertain whether *S* could verbalize the contingencies of *E*'s verbal statement, whether he could relate them to his own performance, and whether he had ever undergone a similar experiment before. *S*s in both groups were further divided into awares (A) or unawares (UA) on the basis of the questionnaire, yielding four subgroups: IH-A, IH-UA, MH-A, and MH-UA.

C. Results

1. *Awareness*

All *S*s who stated that reinforcement was contingent on the use of a hostile response were classified as "awares." The class of hostile responses could be described as "bad words," hostile or aggressive words, words indicating harm or damage, etc. Only descriptions which conveyed the concept of hostility were accepted. In order to test the reliability of the classification three judges who did not know *S*'s group membership, independently sorted a sample of 48 questionnaire verbalizations. Only three replies yielded disagreement among the judges. These *S*s were assigned to the classification given by two of the three judges. The distribution of awareness responses is shown in Table 1. Thirty per cent of all *S*s were classified as aware. A test on the distribution of aware and unaware *S*s among groups MH and IH, yielded an X^2 value of 4.41 ($p < .05$). When corrected for continuity, $X^2 = 3.44$, $p = .07$. Presentation of more easily discriminable word pairs (neutral-intensely hostile) resulted in increased verbalization of the response-reinforcement contingency. These findings support the hypothesis about the relationship between stimulus discriminability and awareness.

TABLE 1
DISTRIBUTION OF AWARENESS AMONG GROUPS

	Aware	Unaware	N
MH	8	33	41
IH	16	23	39
	—	—	—
Total	24	56	80

2. *Pre-acquisition Performance*

The frequency of selecting critical responses in the pre-acquisition block (Trials 1-20) did not differ significantly for Groups IH and MH ($F = 1.07$ with 1 and 76 *df*), nor for *S*s who were later aware *vs.* those who were unaware ($F = 0.54$). The analysis of variance used the correction for disproportionality described by Wert, Neidt, and Ahmann (1954). This finding indicated that neither the discriminability of pairs, i.e., the relative intensity of the hostile member of the word pair, nor whatever tendency in *S*s might differentiate them later with regard to awareness, affected the initial selection of words.

3. *Analysis of Learning Trends*

An unweighted means trend analysis was performed on the mean number of critical responses over all blocks. These data yielded a significant difference in the linear component of the acquisition curves, for aware *vs.* unaware *S*s (Table 2). The linear increment of critical responses over blocks was significantly greater in aware *S*s. The interaction between awareness and stimulus similarity fell just below the 5 per cent level of significance (at $p = .05$, $F = 3.98$ with 1 and 76 *df*). Stimulus discrepancy did not significantly affect learning slopes in their linear component. A t-test for linearity on all unaware *S*s showed that the group slope differed significantly from zero ($t = 3.91$), indicating that these *S*s also showed learning, although at a lower rate than aware *S*s.

TABLE 2
TREND ANALYSIS OF CRITICAL RESPONSES OVER ACQUISITION BLOCKS

Source	df	Linear		Quadratic	
		Ms	F	Ms	F
MH vs. IH (A)	1	.296	2.53	.396	7.92**
Awareness (B)	1	.580	4.96*	.110	< 1
A x B	1	.454	3.88	.042	< 1
Error	76	.117		.050	

*$p < .05$.
**$p < .01$.

The significant difference in the quadratic component between the IH and MH groups suggests the effect of stimulus discriminability on the manner in which learning occurs. A *t*-test for the quadratic component shows no significant component ($t = 1.85$) in group IH. Group MH shows a significant negative quadratic component ($t = 2.14$), indicating an initial steep rise in the curve form with subsequent leveling-off. The trend analysis indicates that learning is better in aware *S*s, regardless of the group in which they occur. However, it also indicates that both aware and unaware *S*s *do* learn and that the change of critical response frequencies over trial blocks differs as a function of stimulus discriminability.

4. *Analysis of Difference Scores*

Earlier research in verbal conditioning has shown considerable variability in the acquisition of this task, both between *S*s and within *S*s over trials. Therefore, separate evidence for learning can be obtained by a measure of change in response frequency from the pre-acquisition to the last (fifth) acquisition block. This measure uses each *S*'s individual pre-acquisition performance as a basis for improvement and also reduces trial-to-trial variability. It differs from the trend analysis in asking only whether groups differ in their *relative* change from the beginning to the end of acquisition.[2]

The mean improvement scores for the groups are presented in Table 3. All groups show some increase over their pre-acquisition performance. A *t*-test was run on each group to determine whether its scores represented significant increases in the use of the reinforced response class (Table 3).

TABLE 3
MEAN IMPROVEMENT SCORES AND T-TESTS FOR ALL GROUPS

	Aware				Unaware			
	N	Mean		SD	N	Mean		SD
MH	8	3.25	t = 2.11*	4.37	33	2.70	t = 5.78**	2.69
IH	16	3.38	t = 3.59**	3.76	23	0.61	t = 0.63	4.60

* $p < .10$.
** $p < .01$.

Groups MH-UA (unaware) and IH-A (aware) clearly revealed significant improvement. Group IH-UA (unaware) showed no change as a result of training. In group MH-A a trend toward learning is noted. While the *t* value falls just short of the value required at $p = .05$, these findings should be considered as highly suggestive of learning in this group, when the small group size and the relatively high mean change in the predicted direction are taken into account. In order to examine the extent of the relationship between learning and awareness, point-biserial *r*'s were run. In MH groups $r_{ps} = 0.074$, $t = 0.46$, $p > .05$, in IH groups $r_{ps} = 0.32$, $t = 2.08$, and $p < .05$. These data reveal a moderate significant correlation between awareness scores and learning scores only for *S*s who were presented with highly distinct stimuli.

Differences in improvement scores among the groups were examined by an analysis of variance, adjusted for cell disproportionality (Table 4). The results indicate that the discriminability of words within stimulus pairs (MH *vs.* IH) did not result in differential improvement during acquisition ($F = 2.89$). However, *S*s who were categorized as aware, regardless of the stimuli presented, showed a significantly greater increase in use of critical responses than unaware *S*s ($F = 4.04$). The significant Hostility x Awareness interaction suggests that the discriminability of stimuli differentially affected aware and unaware *S*s. A simple effects test revealed significantly more improvement in group MH-UA than in IH-UA ($F = 5.57$ with 1 and 54 *df*). The aware *S*s, (MH-A *vs.* IH-A) did not differ significantly from each other.

---

[2] Ferguson and Buss (1960) present a further argument for use of difference scores as the most conservative measure of learning.

TABLE 4
Analysis of Variance of Improvement Scores in All Groups

Source	df	MS (adjusted)	F
MH vs. IH (H)	1	40.45	2.89
Awareness (A)	1	55.43	4.04*
H × A	1	54.65	3.98*
Error	76	13.73	

* $p < .05$.

D. Discussion

The results clearly support the hypothesis that greater distinctiveness between the S^D and S^Δ in a Taffel-type verbal conditioning task results in a higher proportion of Ss who can specify the response-reinforcement contingency or the properties of the S^D. This finding indicates that the probability of obtaining awareness reports in verbal conditioning is at least partially a function of the class of discriminative stimuli and could be systematically varied. In a free operant situation, Matarazzo et al. (1960), studied the relationship between response classes which differed in discriminability (plural nouns versus humans) and found awareness and conditioning only when the response class "humans" was reinforced. Other attempts to vary awareness by instructions are reported by Sidowski (1954) and Krasner, Weiss, and Ullman (1959); by varying subject characteristics in a study by Medini (1957); and by use of different questionnaires in Levin's study (1959).

The relative contributions to the verbal conditioning process by the stimulus variables and the awareness variable have not been completely isolated in this study. On the one hand, both the analyses of trends and difference scores clearly show that aware *S*s in both groups yield higher performance in acquisition than unaware *S*s. In addition, stimulus disparity also leads to a higher incidence of awareness responses. On the other hand, the stimulus similarity variable did not affect learning in the same way as it affected awareness reports. With use of improvement scores the only differential effect of MH *vs.* IH is noted in unaware *S*s. While some "learning without awareness" is observed in the MH-UA group, the IH-UA group clearly fails to learn. These results do not support our hypothesis that an easier discrimination results in better performance. It appears that this hypothesis which is supported so easily by research in most other areas of learning that it is almost axiomatic, must be modified by taking into consideration some other processes conveniently labeled as "awareness" in the verbal conditioning situation. The expected superiority of the IH group not only fails to appear in the aware *S*s but the expectations are reversed in the unaware *S*s. The significant quadratic component in group MH also indicates that stimulus discriminability may result in differential changes from block to block during acquisition. However, the difficulty of interpreting the meaning of this finding and the lack of a specific a priori hypothesis about this component warrants caution in the generalization of these results. In the present study the MH group showed the greatest increment early while the IH group improved at the same rate over all blocks.

It is interesting to consider the possibility that the specific content of the classes used here may have affected the results. Differential conditioning effects with hostile or aggressive responses have been reported by Buss and Durkee (1958) as a function of sex difference. Similarly, Binder, McConnell, and Sjoholm (1957) report differences in performance with hostile words when *E*s of different sex are used, and explain these results in terms of inhibition of hostility in the presence of the male E. Anderson (1958) found that aggressive *S*s showed higher verbal conditioning. If some *S*s in our IH group tended to avoid usage of intensely hostile words, such avoidance might also lead to a lower likelihood of discriminating this stimulus class. In this case, *S*'s behavior may be similar to that described as repression and account for the poor performance in the IH-UA group. That some subject characteristics (personality variables) also affected our findings is demonstrated clearly when the performance of individual *S*s in each group is scrutinized. Although group IH-UA shows a slight mean increase, closer analysis revals that in this group 10 *S*s increased in their responding while 11 *S*s decreased in usage of intensely hostile words, despite reinforcement. Two *S*s showed no change. Thus, almost 50 per cent of this group actually showed a change in the direction of the conditioning effects. In the MH-UA group 30 *S*s showed an increase while only three showed a decrease. In the aware groups, IH-A 10 *S*s showed an increase, two *S*s decreased, and four *S*s failed to change. In the MH-A group seven *S*s increased in usage of the critical class. This analysis clearly indicates that the group findings do not simply represent a difference in the extent to which acquisition was successful, but the group data also mask individual differences in the direction in which the social reinforcer showed its effectiveness in individual *S*s. Of all *S*s, only six remained unaffected by reinforcement, while all others changed in one direction or another. The group findings actually represent the proportion of *S*s which moved in one direction or another during the acquisition as well as the extent of their change. This breakdown emphasizes the differential responsiveness among individuals to social reinforcement in the verbal conditioning situation, and suggests that much of the currently observed variability in these studies may simply represent the effect of variables which have not yet been clearly isolated.

E. SUMMARY

The effect of the disparity between discriminanda on learning during acquisition and on awareness was investigated in a verbal conditioning situation. Eighty *S*s were divided into two experimental groups. On the basis of a post-experimental questionnaire each group was further subdivided into aware and unaware *S*s. In group MH, *S*s were given word pairs consisting of a neutral word and a mildly hostile word. In group IH, *S*s were given pairs consisting of a neutral word and an intensely hostile word. *E* said "good" whenever *S* constructed a sentence using a hostile word. One hundred training trials were given. The results indicated that group IH showed a significantly higher number of aware *S*s than group MH. Aware *Ss* in both groups showed learning during training. Of unaware *S*s only those who were in group MH showed learning. Analyses of learning trends and improvement scores yielded significant differences between aware and unaware *S*s but not between groups MH and IH, except for the quadratic component of the curves. A significant interaction between awareness and groups showed a superiority of unaware *S*s in group MH who learned, over unaware *S*s in group IH, who did not learn.

REFERENCES

1. ANDERSON, D. E. Personality variables and verbal conditioning. *Unpbl. Ph.D. Dissert.*, Univ. Nebraska, 1958.
2. BINDER, A., MCCONNELL, D., & SJOHELM, N. A. Verbal conditioning as a function of experimenter characteristics. *J. Abnorm. & Soc. Psychol.*, 1957, **55**, 309-314.
3. BUSS, A. H., & DURKEE, A. Conditioning of hostile verbalizations in a situation resembling a clinical interview. *J. Consult. Psychol.*, 1958, **22**, 415-418.
4. DULANY, D. E. Hypotheses and habits in verbal "operant conditioning." *Psychol. Rev.* (In press.)
5. FERGUSON, D. C., & BUSS, A. H. Operant conditioning of hostile verbs in relation to experimenter and subject characteristics. *J. Consult. Psychol.*, 1960, **24**, 324-327.

6. Krasner, L. Studies of the conditioning of verbal behavior. *Psychol. Bull.*, 1958, **55**, 148-170.
7. Krasner, L., Weiss, R. L., & Ullman, L. P. Responsivity to verbal conditioning as a function of two different measures of "awareness." Paper read at the A.P.A., Cincinnati, 1959.
8. Levin, S. The effects of awareness on verbal conditioning. *Unpbl. Ph.D. Dissert.*, Duke Univ., 1959.
9. Matarazzo, J. D., Saslow, G., & Pareis, E. N. Verbal conditioning of two response classes: Some methodological considerations. *J. Abnorm. & Soc. Psychol.*, 1960, **61**, 190-206.
10. Medini, S. Learning without awareness and its relationship to insight and the hysteric-obsessive dimension. *Unpbl. Ph.D. Dissert.*, New York Univ., 1957.
11. Sidowski, J. B. Influence of awareness of reinforcement on verbal conditioning. *J. Exper. Psychol.*, 1954, **48**, 355-360.
12. Taffel, C. Anxiety and the conditioning of verbal behavior. *J. Abnorm. & Soc. Psychol.*, 1955, **51**, 496-501.
13. Wert, J. E., Neidt, C. O., & Ahmann, J. S. Statistical Methods in Educational and Psychological Research. New York: Appleton-Century, 1954.

Department of Psychology
Purdue University
Lafayette, Indiana

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 125-131.

TRENDS TOWARD MULTIPLE AUTHORSHIP IN SCIENTIFIC PUBLICATIONS*

Bionics and Computer Branch, Electronic Technology Laboratory
Wright-Patterson Air Force Base

MILDRED B. MITCHELL

A. INTRODUCTION

Mapheus Smith (17) found a trend toward multiple authorships in papers read at the American Psychological Association from 1946-1957. He described the trend as being much stronger in some divisions of APA than in others. Smith suggested that one reason for the trend toward multiple authorships, and especially for the relatively large number of multiple authors in the Military Division, was the fact that many psychologists spent the war years as members of large staffs of government agencies working on projects of great urgency. Now many are continuing the work pattern, with less sense of urgency, in government agencies, in private institutions or large state universities supported by government contracts requiring several psychologists or other scientists (3, 5). Government support of research has not only increased in this country but is prevalent in other countries as well (8, 10, 13, 14, 16).

B. METHOD

Smith (17) further suggested that a study similar to his should be made for listings in *Psychological Abstracts* and non-psychological literature. This appealed to the writer as a simple, straightforward procedure which could be pursued independently while working on a large military project which could not be handled independently. When an interest in comparing American with non-American psychological abstracts developed, however, it turned out to be not so simple a task. It led to tracking down the place of publication of hundreds of journals in 50 different countries (6, 7, 11, 12). The task was complicated by the fact that some of the journals had changed their names several times. In contrast to the total of 4,189 psychological papers Smith tabulated for number of authors for 12 years, more than 100,000 entries were tabulated for 15 years in the *Psychological*

---

* Received in the Editorial Office on March 8, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

Abstracts and considerably more than 100,000 entries were tabulated for two years in the *Chemical Abstracts.*

C. RESULTS

A tabulation of the percentage of papers by single and multiple authors in *Psychological Abstracts* for the years 1946-1958 showed a trend similar to that found by Smith in papers read at APA meetings in 1946-1957, but it was much less marked (Figure 1). Possibly the policy of many employers to finance trips to psychological meetings *only if* the employee's name appears on the program has some influence on this discrepancy (2).

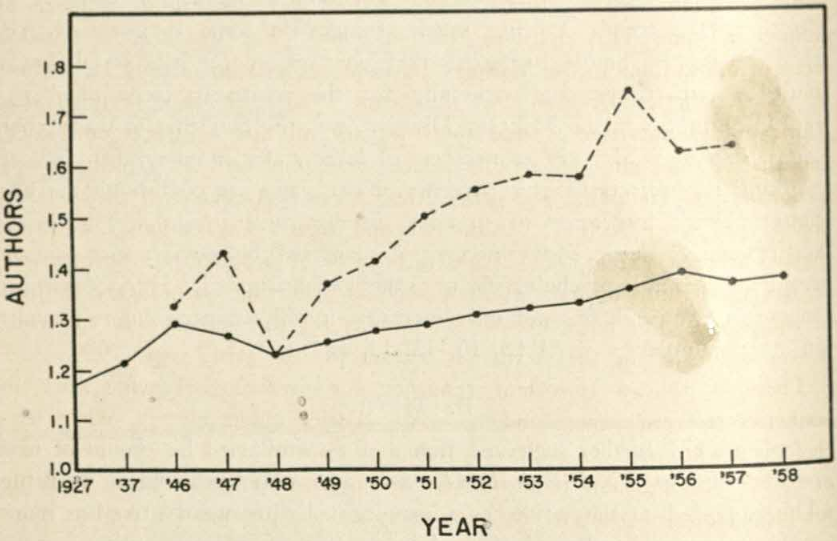


FIGURE 1
MEAN AUTHORS APA PAPER FROM SMITH (17) 1946-1957 - - - - -
MEAN AUTHORS PER PSYCHOLOGICAL ABSTRACT 1927-1958 ————

Since the trend toward an increase in multiple authorship in *Psychological Abstracts* for 1946-1957 was so much lower than for the APA papers for the same years, it was decided to look for a trend in the abstracts over a longer period. A tabulation was, therefore, made for one year in each decade beginning with the first volume of *Psychological Abstracts,* 1927. There is a fairly consistent decrease in percentage of single author abstracts with a corresponding increase in percentage of two-author, three-author, and four-or-more-author abstracts for the years 1927, 1937, 1947 and 1957 (Table 1).

TABLE 1
PERCENTAGE OF SINGLE AND MULTIPLE AUTHORSHIPS OF MATERIAL ABSTRACTED IN PSYCHOLOGICAL ABSTRACTS

	Number of authors				
Year	1	2	3	4+	N
1927	85.7	12.3	1.4	0.6	2,674
1937	83.2	13.9	2.5	0.5	5,913
1947	80.1	14.7	4.1	1.1	4,518
1957	74.4	17.5	6.1	2.1	8,951
Total	79.3	15.3	4.1	1.3	22,056

As mentioned above, Smith (17) found a higher mean number of authors in some APA divisions than in others. For instance, he found a mean of 1.8 authors in the Military Division but a mean of only 1.2 authors in the Aesthetics Division for the years 1946-1957. The *Psychological Abstracts* did not have a separate listing for military abstracts until 1959 and then it had only 16 articles which were published as a military subsection under "Industrial and other Applications." A review of the abstracts under the entire section, "Industrial and other Applications," however, showed at least 152 of the 375 abstracts to be of military technical reports, or having definite military content such as reporting on military subjects or equipment. The mean number of authors for this section was 1.6 while it was only 1.0 for the Aesthetics section for the same year, 1959.

There is thus a consistent tendency for psychologists who work in aesthetics to read papers and to write articles independently, while psychologists who work in military psychology tend to work in groups of two or more. This probably is due in part to government money being available for large military projects requiring complicated equipment as well as more personnel.

Agency policies regarding paying expenses of employees listed an an author of a paper to be read at APA may, of course, account for some of the difference between the percentage of multiple authors of papers and the percentage of multiple authors of abstracted material but another possible cause for the discrepancy must be considered. That is the country of origin. Almost all participants in APA programs are Americans. On the other hand, although the percentage of abstracted material originating outside the USA has been influenced by geopolitics, the percentage has been considerable from the beginning. For the years 1927, 1937, 1947 and 1957 it was 34 per cent, 45 per cent, 19 per cent and 26 per cent respectively.

A tabulation of single and multiple authorships by country of publication for the years 1927, 1937, 1947, and 1957 gives some evidence to support the hypothesis that the inclusion of articles from other countries may tend to lower the percentage of multiple authorships. Since the number of abstracts vary greatly from country to country and from year to year, especially during the war years, the results are shown only for the six individual countries which have a total of more than 250 abstracts for the four years studied (Table 2). Each of these six countries, namely, France, Germany, Great Britain, Italy, Russia, and Switzerland, as well as the Scandinavian countries and the total for Other Countries had a higher percentage of single authors than did the USA. In fact, the percentage of non-USA (86.4) single authors for these four years was more than 10 per cent greater than that for the USA (76.1). The range of percentages for single authorships outside the USA was from 81.0 for Great Britain and Russia to 91.8 for Germany. Switzerland ran a close second with 91.3 per cent single authorships. A slight trend for increase in multiple authorships is seen outside the USA, especially for 1957, but it is not as marked nor as consistent as in the USA (Table 3). The drop in percentage of single authorships for non-USA was from 87.7 to 82.8 for 1947 to 1957. The foreign authors, however, do not account for all the differences in multiple authorships between papers read at APA and the *Psychological Abstracts*. Smith (17) reported a drop in percentage of single authors of papers from 70 per cent in 1947 to 52 per cent in 1957. The drop for American abstracts was only from 78 per cent in 1947 to 71 per cent in 1957 (Table 3).

TABLE 2

PERCENTAGE OF SINGLE AND MULTIPLE AUTHORS ABSTRACTED IN PSYCHOLOGICAL ABSTRACTS BY COUNTRIES FOR TOTAL OF YEARS 1927, 1937, 1947, & 1957

	Number of authors				
Country	1	2	3	4+	N
Germany	91.8	7.0	1.0	0.2	1,645
Switzerland	91.3	7.1	1.1	0.5	367
Italy	88.3	10.5	0.9	0.3	334
Scandinavian	88.0	9.8	2.3	0.0	266
France	81.7	12.0	4.3	2.0	887
Great Britain	81.0	14.5	3.0	1.4	1,286
Russia	81.0	13.7	3.5	1.8	342
Other non-USA	87.2	10.1	2.2	0.5	1,654
All non-USA	86.4	10.5	2.3	0.8	6,781
USA	76.1	17.5	5.0	1.5	15,275
Total	79.3	15.3	4.1	1.3	22,056

TABLE 3
PERCENTAGE OF SINGLE AND MULTIPLE AUTHORS OF MATERIAL ABSTRACTED IN PSYCHOLOGICAL ABSTRACTS FROM USA AND NON-USA, 1927, 1937, 1947, AND 1957

	Number of authors									
	Non-USA					USA				
Year	1	2	3	4+	N	1	2	3	4+	N
1927	87.7	10.0	1.6	0.6	937	84.6	13.6	1.3	0.5	1,737
1937	88.8	9.2	1.4	0.6	2,651	78.5	17.7	3.3	0.5	3,262
1947	87.7	9.9	1.9	0.6	843	78.3	15.8	4.7	1.2	3,675
1957	82.8	12.3	3.7	1.2	2,350	71.3	19.3	6.9	2.4	6,601
Total	86.4	10.5	2.3	0.8	6,781	76.1	17.5	5.0	1.5	15,275

Perhaps government support of research in other countries has also influenced the percentage of multiple authorships. We do not have information regarding government support of research in all the countries considered, but we do have some information regarding Great Britain and Russia, the two countries closest to the USA in percentage of multiple authorships. Hiscocks (8) estimates that 75 per cent of research in the United Kingdom is government supported. In Russia, of course, all research is government supported (13, 16).

As Smith (17) had suggested, it seemed of interest to compare single and multiple authorships of psychologists with that of other scientists. Since *Chemical Abstracts* covers a wide range of scientific interests in addition to chemistry, the entries for years 1927 and 1957 were tabulated (Table 4). In the first place, we might say that the same trend holds for the *Chemical Abstracts* as for the *Psychological Abstracts,* only more so. The percentage of single authorships tends to decrease and multiple authorships to increase. The mean number of authors in *Psychological Abstracts* increased from 1.2 to 1.4 while the mean number of authors in *Chemical Abstracts* increased from 1.4 to 1.8 Thus we see that the other scientists represented in the *Chemical Abstracts* seem to work, or

TABLE 4
PERCENTAGE OF SINGLE AND MULTIPLE AUTHORS IN PSYCHOLOGICAL ABSTRACTS AND CHEMICAL ABSTRACTS 1927 AND 1957

	Number of authors											
	Psychological abstracts						Chemical abstracts					
Year	1	2	3	4+	Mean	N	1	2	3	4+	Mean	N
1927	85.7	12.3	1.4	0.6	1.2	2,674	70.7	24.7	3.8	0.8	1.4	31,088
1957	74.4	17.5	6.1	2.1	1.4	8,951	47.7	33.6	13.1	5.6	1.8	101,452

at least publish in groups more frequently than psychologists. The difference appears even more remarkable when we note that both the *Psychological Abstracts* and *Chemical Abstracts* cover some of the same journals. The mean number of authors for *Chemical Abstracts* was the same in 1927 (1.4) as it was for *Psychological Abstracts* in 1957. Furthermore, the percentage of single authors in *Psychological Abstracts* was greater in 1957 (74.4) than the single authors in *Chemical Abstracts* in 1927 (70.7).

The data analyzed seemed to indicate a definite trend toward multiple authorship by psychologists and by other scientists. It should be pointed out that the data presented are not entirely complete, but if they were, they would only emphasize the trend because abstracts listed with *et al.* were omitted. Smith (17) noted four papers in the post-war year, 1946, with a vague "staff" authorship. The *Psychological Abstracts* also carried a few such abstracts, especially reports of military efforts, probably vague because of the rapidly changing military personnel. The largest source of error in both *Psychological Abstracts* and *Chemical Abstracts,* however, was in the use of *et al.* after 1, 2, 3, and 4 names. Obviously these are multiple authorships, but how many? Since *Psychological Abstracts* listed as many as 15 authors and *Chemical Abstracts* listed as many as 22 authors, it first seemed sensible to include all *et al.* as four-or-more-authors. After checking a small sample, however, it was found that some authors or abstractors used the *et al.* when only three authors were involved. There apparently had been no consistent editorial policy or enforcement of editorial policy regarding the use of *et al.,* so it was decided simply to eliminate it.

D. Summary and Conclusions

Undoubtedly, Smith is correct in crediting government financing of research, both within its own agencies and its contracts with universities and industries, for some of the increase in multiple authorships of psychological and other scientific publications. Perhaps also the relatively young science of Psychology with its recently formulated Code of Ethics (1) is becoming more conscious of the policy, so strongly advocated and practiced by Hull (9), of allowing students and assistants to receive credit for their work instead of having everything published in the professor's or laboratory chief's name. We have no data to check these hypotheses, but we have presented results including a very large number of abstracts to suggest that there has been a definite trend toward multiple authorships and a decrease in single authorships both for publications covered in the *Psychological Abstracts* and in the *Chemical Abstracts* from 1927 to 1957. The discrepancy between the percentage of multiple authorships in papers read at the American Psychological Association meetings and in articles abstracted in the *Psychological Abstracts* can only partially be accounted for by the fact that there are a relatively large number of abstracts from other countries while most of the papers are given by Americans. The policy of many agencies to pay expenses to conventions only if the employee's name appears on the program probably also accounts for some of the discrepancy.

REFERENCES

1. AMERICAN PSYCHOLOGICAL ASSOCIATION. Ethical standards of psychologists. Report. *Amer. Psychol.*, 1959, **14**, 279-282.
2. AMERICAN PSYCHOLOGICAL ASSOCIATION. Technical communication in psychology: A statement of the problem. *Amer. Psychol.*, 1959, **14**, 267-271.
3. BERKNER, L. V. Government sponsorship of scientific research. *Science*, 1959, **129**, 817-821.
4. Chemical Abstracts. Vol. 21 & 51. Columbus, O.: American Chemical Society. 1927 & 1957. Pp. 5512 & 18,702+ 5682 (Indices).
5. DARLEY, J. G. The impact of federal support on higher education. *Amer. Psychol.*, 1959, **14**, 480-488.
6. FRANCK, M. (Ed.) Union List of Serials in Libraries of the United States and Canada. Supp. II. Jan. 1944-Dec. 1949. New York: Wilson, 1953. Pp. 1365.
7. GREGORY, W. (Ed.) Union List of Serials in Libraries of the United States and Canada. New York: Wilson, 1943. Pp. 3065.
8. HISCOCKS, E. S. Organization of science in the United Kingdom. *Science*, 1959, **129**, 689-693.
9. HULL, C. L. *In* Boring, E. G., Langfeld, H. S., Werner, H., and Yerkes, R. M. (Eds.) History of Psychology in Autobiography IV. Worcester, Mass.: Clark U., 1952. Pp. 143-162.
10. MAJOR, R. Organization of scientific activities in Norway. *Science*, 1959, **129**, 694-700.
11. MALIKOFF, G. W. (*Ed.*) Union List of Serials in Libraries of the United States and Canada. Supp. I. Jan. 1941-Dec. 1943. New York: Wilson, 1945. Pp. 1123.
12. MORTON, L. T. World Medical Periodicals. (2nd. Ed.) New York: World Medical Assoc., 1957. Pp. 340.
13. MURRAY, H. A., MAY, M. A., & CANTRIL, H. Some glimpses of soviet psychology. *Amer. Psychol.*, 1959, **14**, 303-307.
14. PRICE, D. K. Organization of science here and abroad. *Science*, 1959, **129**, 759-765.
15. Psychological Abstracts. Vol. I, 11, 20-33. Washington, D.C.: Amer. Psych. Assoc., 1927, 1937, 1946-1957. Pp. 722, 735, 585, 620, 774, 886, 804, 922, 844, 921, 952, 931, 866 & 875.
16. RAZRAN, G. Soviet psychology and psychophysiology. *Science*, 1958, **128**, 1187-1194.
17. SMITH, M. The trend toward multiple authorship in psychology. *Amer. Psychol.*, 1958, **13**, 596-599.

1726 Kensington Drive
Dayton 6, Ohio

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 133-136.

HOW FAT IS FAT? SOME CONSEQUENCES OF SIMILARITY BETWEEN JUDGE AND JUDGED OBJECT*

Department of Psychology, University of North Carolina

SAMUEL FILLENBAUM

A. INTRODUCTION

This study is an attempt to interpret the relation between judgment on an attribute and own standing on that attribute in terms of the similarity of the object of judgment to the judge. Considerable evidence suggests the importance of a reference or comparison level in the making of judgments (Helson, 1959). Further, it has been found that own standing on an attribute may function as such a comparison level, the scale of judgment being organized around, or at least displaced with regard to, one's own position. Thus in an early study (Marks, 1943) a significant relation was found between rater's skin color and skin color ratings made for other *S*s, the darker *Ss* rating a given object person lighter than the lighter colored *Ss*. Findings with similar implication have been reported with regards to estimates of height, taller men were found to center their scales around a significantly higher mean than shorter men (Hinckley and Rethlingshafer, 1951). An additional study where age was one of the variables, and own age was considered as the background or residual stimulus yielded results consistent with the above (Rethlingshafer and Hinckley, 1954). Some psychophysical studies indicate that an item contributes to the reference level with regard to which an object is judged to the extent that the item is seen as similar to the object of judgment. Thus, for example, it has been found that a weight in the form of a box, has significantly less effect on judgment than a weight of the same heaviness which has the same appearance as the object being judged (Brown, 1953).

The above leads to the suspicion that own standing will influence judgment insofar as the judged object is seen as comparable to self with regard to the dimension of judgment. Or to put it more generally, the relation between own standing and judgment of others with regard to some attribute will depend on the similarity between the object of judgment and the judge. This hypothesis was tested by requiring *S*s to make height and

---

* Received in the Editorial Office on March 16, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

weight estimates for same and opposite sex persons. By using both male and female *Ss* and requiring judgments for both sexes an efficient replication of the study was possible, and by using two different dimensions (one previously studied in this context, one not) it was hoped that at least some idea as to the generality of the findings might be obtained.

B. METHOD

The *Ss* were 191 students in the elementary course in psychology at the University of North Carolina; there were 114 male and 77 female *Ss* and they were run in classroom groups ranging in size from 15 to 60. *Ss* were given 5″ x 3″ cards and were asked to put down their estimates of: the height in feet and inches of an average man, an average girl; the weight in pounds of an average man, an average girl. After this had been done they turned the cards over and were then asked to put down their own height and weight, and M or F depending upon their sex.

C. RESULTS AND DISCUSSION

It can be seen from Table 1 that in three of four instances the correlation between own standing and same sex estimates is significant, and greater than the correlation between own standing and other sex estimates. In the fourth case, height ratings by women, while the correlation between own standing and same sex estimates is slightly larger than that between own standing and other sex estimates it is apparent that this difference is negligible and that neither coefficient approaches significance.

TABLE 1
CORRELATION (+) BETWEEN OWN STANDING AND ESTIMATES OF AVERAGE HEIGHT AND WEIGHT

	Men (N = 114) Own standing X		Women (N = 77) Own standing X	
	Same sex estimate	Opposite sex estimate	Same sex estimate	Opposite sex estimate
Height	.47**	.17	.13	.08
Weight	.36**	—.08	.44**	.22

** p <.01 (two tailed).

It was assumed that the relation between own standing and judgment for some attribute would depend on the similarity between judge and judged object, and that consequently there should be a significant relation between own standing and same sex ratings and a smaller or no relation between own standing and other sex ratings. To a considerable extent the results are consistent with hypothesis, seven of the eight correlation coefficients and consequently three of the four critical comparisons fall as expected. For women raters, the relation between own height and same sex estimates of average height presents the one exception, this relation being zero order ($r = .13$). This last finding is difficult to interpret. In some later work which in part replicates this study (Fillenbaum, 1961) for women raters a significant correlation *was* found between own height and estimates of average height for women, this correlation being greater than that between own height and estimates of average height for opposite sex (men).

Considering the relation between own standing and object judgment some general comments might, perhaps, be made. It may be possible to provide a partial interpretation of projection within the general framework of a judgment approach (see Berkowitz, 1960, where such an analysis is worked out in some detail). Given the denial of a certain attribute in the self there should be a shift in reference level for that attribute and it should follow that others would be rated rather higher for that attribute. More specifically, in the light of our results, it might be expected that there would be more projection of the attribute on those individuals who, in other respects, are seen as similar to the self, than on those seen as dissimilar.

The results of this study again suggest, in a context somewhat different from that of the usual psychophysical experiment, that an item will enter in the determination of the reference level with regard to which some object is judged only to the extent that it is seen as comparable or similar to that object. For response prediction knowledge merely of the value of the item on the attribute to be judged is insufficient.

D. SUMMARY

The relation between object judgment for some attribute and own standing on that attribute was examined in terms of the comparability or similarity of the objects of judgments to the judge. *S*s, 191 university students of whom 114 were male and 77 female, were required to make estimates as to average height and weight for same sex and opposite sex persons, and were then asked for their own height and weight. In general, a significant positive relation was found between own standing and same sex ratings, with lesser or no relation between own standing and opposite sex ratings. Of eight correlations seven took values as expected and consequently three of the four critical comparisons were consistent with hypothesis. For women raters, the zero order relation between own height and same sex ratings of average height constituted the one exception.

It was indicated how some aspects of projection might be thought of in terms of a judgment approach, and that something might be said as to the selectivity of projection in terms of such an approach.

REFERENCES

1. BERKOWITZ, L. The judgmental process in personality functioning. *Psychol. Rev.*, 1960, **67**, 130-142.
2. BROWN, D. R. Stimulus similarity and the anchoring of subjective scales. *Amer. J. Psychol.*, 1953, **66**, 199-214.
3. FILLENBAUM, S. Own position in relation to estimates of average standing and desirability for more and less self-comparable objects. *J. Personal.*, in press.
4. HELSON, H. Adaptation level theory. (Pp. 565-621 in *Psychology: A Study of a Science, Vol. I,* S. Koch, *Ed.*). New York: McGraw-Hill, 1959.
5. HINCKLEY, E. D., & RETHLINGSHAFER, D. Value judgments of heights of men by college students. *J. of Psychol.*, 1951, **31**, 257-262.
6. MARKS, E. S. Skin color judgments of Negro college students. *J. Abn. Soc. Psychol.*, 1943, **38**, 370-376.
7. RETHLINGSHAFER, D., & HINCKLEY, E. D. Influence of characteristics of judges on their psychophysical judgments. *Amer. Psychol.*, 1954, **9**, 454.

Department of Psychology
University of North Carolina
Chapel Hill, North Carolina

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 137-147.

THE SENSORY PARALLEL OF THE REORGANIZATION PERIOD IN THE CORTICAL RESPONSE IN INTERMITTENT RETINAL STIMULATION*[1]

Department of Psychology, Michigan State University

S. Howard Bartley, Thomas M. Nelson, and Jane E. Ranney

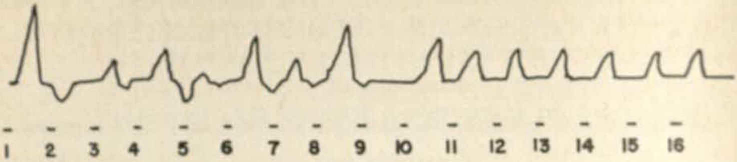
A. Problem

The study of the activity of the optic pathway has largely been conducted by use of three temporal forms of stimulation, namely, by isolated, paired, and indefinitely repeated pulses, the latter often called intermittent stimulation. The same forms of stimulation have been used to study sensory end-results. Bishop and Bartley in a series of papers (1, 2, 3, 4, 9) reported upon an analysis of the activity of the optic pathway, disclosing that any single channel in the pathway from retina to cortex cannot be reactivated at extremely short intervals, but that as a total the pathway may be kept in continuous and uniform activity. From these and a number of associated facts, Bartley (4, 5) put forth his alternation of response theory. This was not a theory or description of how the neuroretina acts, but rather how the central end of the pathway utilizes the inputs given it. Bartley (1) showed, for example, that if a train of photic pulses was delivered to the eye at rates above the rate at which any single channel in the pathway could react repeatedly, cortical reaction to the input was at first irregular. This irregularity was an example of what might be expected in accord with the alternation of response theory. After a few pulses of the train were delivered, this irregularity, including a failure to respond at all to some of the pulses, gave way to uniformity of response, each photic pulse being responded to and with the same sized wave in the record (see Figure 1). The period of irregular response was called the reorganization period. Such observations form the basis of what to expect in comparing the critical flicker frequency (CFF—intermittent rate just producing steady brightness) obtaining after the observer views a target for some time, and the CFF when only a very few photic pulses are delivered. It

---

* Received in the Editorial Office on March 22, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The present study is a part of the work being done under Research Grant NSF-G-5821 from the National Science Foundation.



REORGANIZATION OF CORTICAL RESPONSE AT BEGINNING OF STIMULUS TRAIN

FIGURE 1

A record of the cortical response to stimulation of the retina by a train of photic pulses at a rate above which any given channel in the optic pathway can respond to successive pulses. At first this is evidenced by absence of certain waves in the record and irregular sizes of others. This initial period of irregularity has been interpreted by Bartley from this and other evidence to be a reorganization period. Following the reorganization period, alternation of responses of various channels to successive pulses is regularly established.

would be expected that the temporal separation between pulses for just producing "fusion," i. e., steady brightness, would be considerably greater in the latter case.

Wilkinson (16, 17), in our laboratory, was led to test for certain parallels between the neurophysiological findings just described and the sensory end results from using two to five photic pulses in a train. He found a failure of certain pulses in the trains to produce flashes. This was in line with expectations. Incidentally, he also found that longer temporal separations between pulses were required when only a few pulses were used than when the usual indefinite trains of pulses are used.

Basler (8) earlier reported a longer pulse interval for producing fusion with two pulses than required for indefinitely long trains. He tried unsuccessfully to explain the results by a persistence-of-vision theory.

Aside from expecting that the interval needed to just resolve two pulses in the first part of a train ought to be greater than after the train has continued awhile, it ought to be expected that interval should vary in keeping with the number of pulses in a short train. If only two pulses are presented, considerable separation ought to be required, according to the implications of the alternation of response theory. If progressively more pulses are used, the chances of the pulses used activating some channels is increased, and furthermore, with longer trains, it will be only the latter parts of them that will be discriminated, as will be made clear in the following paragraphs.

A number of considerations are pertinent both in setting up expectations and in interpreting findings once they are obtained. These have to do with realizing what it is that is being compared in the various situations. This is based upon some initial sophistication in regard to the nature of sensory response in situations of this general sort.

In dealing with long trains of pulses, as in the usual flicker experiments, various sorts of visual resolutions can be made. The most common is measured by CFF. It is the determination of the conditions giving, as it were, the boundary between flicker and a steady field. Flicker emerges and becomes coarser and more pronounced (amplitudinous) as pulse rate declines below this boundary point until the observer is able to see complete darkness between flashes. In this range, various phenomenal entities emerge, which we need not deal with here. That photic pulse rate is a crucial factor in bringing about phenomenal differences in the degrees or kinds of resolution that can be made visually must be kept in mind. It must also be kept in mind that each of these resolutions can be taken as an indicator of critical separation between pulses. CFF is the measure of only one form of resolution, the determination of the boundary between flicker and steady brightness. It could be that to deal with resolutions when using a very few pulses is different than when dealing with indefinitely long trains of pulses.

The present experimentation is based on the broad general laboratory experience that very brief and transient phenomena do not yield to very much internal analysis. If one wishes to know what is produced by the first and second pulses in a train, one must limit the presentation to two pulses. If he wishes to know what happens by adding a pulse to the first two, he presents three pulses. This procedure breaks down in certain respects as more and more pulses are added to the trains, for soon the only sensory response possible is the discrimination of what is being produced by the tail end of the train. The discrimination of the first part of the train becomes impossible. This is to say, the observer is able to report only upon the effects produced during the last end of any extended presentation.

When the observational task is detecting flicker produced by simply two, three or some very small number of pulses, the phenomenology may be somewhat different when two pulses are used than when, let's say, six to eight pulses are used. We do not know how abruptly we can pass from seeing a single unitary flash produced by two pulses to seeing some indication of twoness. Whether this minimal recognition of twoness is equivalent in kind to just seeing flicker with long trains of pulses is a question. It doesn't matter if all we are asking about is some form of minimal resolution regardless of any possible phenomenal differences in various cases.

Once the number of pulses exceeds two, the minimal resolution, as we may call it, may begin to take on new characteristics—it might be seen as a ripple (something more like flicker). Furthermore, from Bartley's (1, 2) findings we might possibly expect a difference between the phenomenal resolution between pulses one and two, and pulses two and three. This is actually what Wilkinson found. If with three pulses only one of the two intervals is resolved, shall it be what is desired for a comparison with the results from other numbers of pulses? The same principle may be involved as we go from three pulses to the use of four. This, as has already been said, is the reason for using two, three, four, etc., pulses as separate presentation series rather than simply having the observer report upon the very first effects at the instant an unlimited train is first presented.

In Wilkinson's study (16), when five pulses were delivered, the observer often reported four flashes. Obviously when repetition rates become high in any situation whatsoever, the observer is less and less able to count, or at least to do so meaningfully. Forsyth and Chapanis (10) conducted a study which bore on this matter. While their problem was very different from either Wilkinson's or ours, their findings do give an idea when to expect the counting process to break down and then become useless for such a purpose as in making the kinds of sensory-neurophysiological comparison interesting us here.

These authors studied the ability of human subjects to count pulses (or, shall we say, flashes) when the numbers presented lay within the range of 1 to 20 and when the delivery rates lay within the range of 2.5 to 30 per second. Among other things, they noted that the variability of responses increased as the number of pulses increased.

Taubman (15) studied the ability to count accurately 1 to 10 pulses of .04 sec. duration when delivered at one or the other of five intervals, .500, .333, .250, .200, and .143 sec. There was, of course, a tendency to underestimate the number delivered. This underestimation was zero when the interval was .500 sec., and the responses were irregular in their accuracy with intervals of .250 or less. At .333 sec. the number of pulses judged was three or less.

Gebhard and Mowbray and colleagues (11, 12, 13) have studied various features of the response of the visual system to temporal variations. These should be consulted by those interested in this general subject and the comparisons between visual and auditory mechanisms, but they are not close enough to the specific purpose at hand to discuss here. The same may be said of the work of Gerathewohl (14).

B. APPARATUS AND PROCEDURE

The instrumentation used to provide short trains of pulses consisted of a large revolving disk (4 feet in diameter) interrupting a beam from a projector. The slide holder of the projector contained an opal glass plate so that the observer looking toward the projector saw a uniform circular luminous target, ⅝ inch in diameter. A portion of the periphery of the revolving disk was sectored into 12 sectors. For one part of the study the separations between these open sectors was one-third the subtense of the sectors, so that the on-off cycle of stimulation, (PCF, pulse-to-cycle fraction) was ¾. In a second part of the study, PCF was changed to ⅛ by use of another disk. For each revolution of the disk, when PCF was ¾, the time required for each open sector to pass a point was 23.2 msec., and for the solid sectors, 5.8 msec., when the revolution rate of the disk was 1 per sec. Since the width of the open sector for PCF ⅛ was about the same as the ⅝ inch field of the luminous target, the duration of the photic pulse at full intensity was unextended. That is, as soon as it reached full intensity, it began quickly to wane to zero. When PCF was ⅛, the open sector passed a given point in 3.6 msec., when the rate was 1 per sec. The stimulus trains were, of course, varied in the number of pulses. To accomplish this, the open sectors not needed were covered by black cardboard, held on by clips. Thus it was only necessary to remove or replace the cardboards from trial to trial to change the number of pulses used.

The rate of the revolving disk and thus the duration of the pulses and intervals between was regulated by varying the speed of the motor driving the disk. This rate was controlled by the experimenter.

Since the photic beam passed by the open sectors was not pointlike but was ⅝ inch in diameter, the onset and decline tapers of the photic impingement were determined by the time required for the disk perimeter to move through an arc of ⅝ inch. This was calculated to be about 2.2 msec. when the revolution rate was 1 per second.

Seven observers in addition to the two senior authors were used for pulse trains from 2 to 6. For the two senior authors, the range of pulse train was from 2 to 12, ordinarily using only the even number of pulses beyond six. Later, the authors obtained the usual CFF's for PCF's of ¾ and ⅛ by observing when the entire peripheries of the disks were sectored. The task of the observers in all cases was simply to indicate when flicker was just eliminated as pulse rate was increased and indicate when it just appeared as pulse rate decreased. Two ascending and two descending series were made for each of the observers, for each of the required pulse trains for each of three pulse intensities, namely, 100.1 c/ft.^{2}, 146.8 c/ft.^{2} and 293.7 c/ft.^{2}

C. Results

The results are shown in Figures 2 to 4. Figure 2 contains the graphs for the combined averages of the two authors for PCF's of ¾ and ⅛. It will be seen that as number of pulses in the train is increased the pulse rate to produce fusion (steady brightness) rises. At first this rise is sharp.

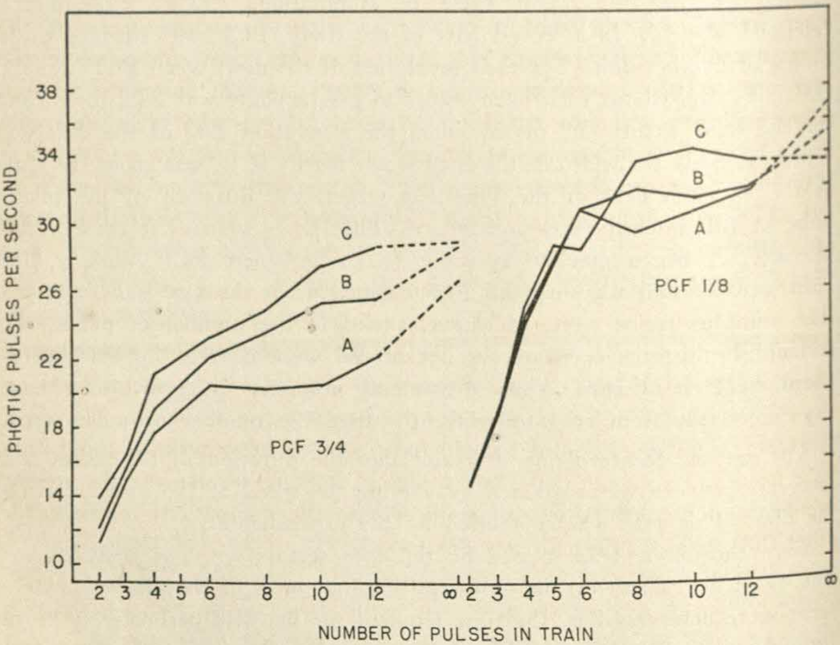


FIGURE 2

The relation between number of photic pulses in a train and the rate at which the pulses must be delivered to produce "fusion" (just eliminating "flicker"). The curves shown are the means of the combined data of the two senior authors. Curves A are for the high intensity used; B, for the medium, and C, for the low intensity. The values at the extreme right are for trains of indefinite length, hence the values are the usual CFF's.

The curves for all three intensity conditions inflect at about six pulses. Subjectively speaking, there is a difference between what is seen when a few pulses are delivered and when the train is longer. With a few pulses, the primary feature ("flicker") is a kind or undulation of flash. With a greater number, the salient feature of each presentation becomes more pronouncedly an extended duration of light with ripples (flicker) superimposed upon it. The inflections in the curves and the region of transition between the two sorts of appearances are concomitant.

It will be noted that the curves for both PCF's ¾ and ⅛ are quite regular for about the first six pulses and then irregularity sets in. This is not at all surprising when the nature of the task is considered. The subjects themselves do not control the rate of increase or decrease in pulse frequency. The subjects simply view the target as fixedly as possible, and if for any reason there is a lapse of attention, an eye blink or any other irregularity, the subject may be a little uncertain where in the observation to signal fusion. He may, therefore, wait and observe the next one or two presentations, and by that time the intervals between pulses will have been increased or decreased beyond what they otherwise would be. This introduces a wider range of variability into the readings than is present when the observer can manipulate the presentation features himself. The momentum of the large disk precludes ready reversals in rate of revolution.

In connection with the fact just mentioned, it must also be realized that there is flicker in the field surrounding the target, and this is very difficult to distinguish from the flicker within the target, especially at the intensity levels used in the experiment. This, too, is a factor working toward variability or readings when the longer pulse trains are delivered. A greater number of readings could have been taken, but these would only have produced a smooth curve by averaging. As it is, we can use the present number and not attribute any significance to the fluctuations in the curve other than what we have already discussed.

It will be noted that the intensity levels were sufficient to separate the three curves for the PCF ¾ throughout their entire length. Not so for the curves for PCF ⅛. The curves for PCF ⅛ cross each other, and this may turn out to have some significance when more analysis of such matters is undertaken.

Figures 3 and 4 are the individual curves for the seven additional observers, who, of course, did not observe for the higher number of pulse trains nor for the extended intermittencies. Considerable divergence in actual values is shown in the curves for both PFC ¾ and ⅛. There was considerable consistency in the relative positions of the curves for the separate observers under all conditions, with the exception of those for M, who possesses some color deficiency ("color blindness"). Whereas he was consistently low for a PCF of ¾, he was just as consistently high for a PCF of ⅛. This would, to say the least, appear to be an example of the phenomenon we have studied and reported elsewhere (6) in which timing and PCF changed the color end results. Conversely, one might expect individuals with color deficiencies to be shifted in their behavior here. The studies just referred to were related to the neurophysiological studies which indicate the existence of a difference in latency, etc., of the responding neural elements to various parts of the spectrum.

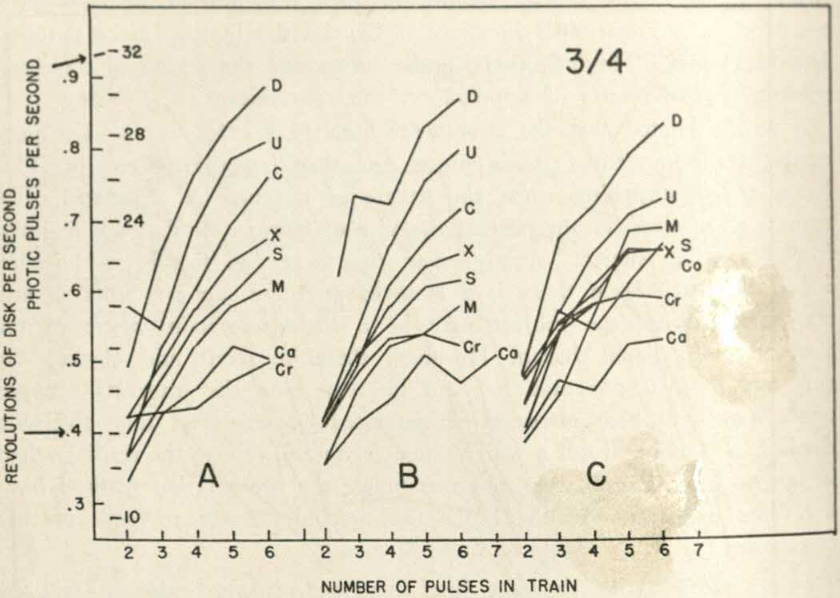


FIGURE 3

The relation between number of photic pulses in a train and the rate for producing "fusion." Each curve in each family is for a separate observer. Family A is for high; B, for medium, and C, for low intensity. The PCF is ¾. Curves X are means of the families.

D. Discussion

First of all, our results were in line with expectations derived from the original neurophysiological investigations of Bartley (1, 2) and with the findings of Wilkinson (16, 17), and thus are in line with the alternation of response theory.

It is obvious that the human observer is not very able to report differences in temporal separation between the various pulses within the trains. He simply reports on numbers of flashes seen when separations are great enough, or upon overall steadiness or lack of flicker. Taking the present work and the work of Wilkinson together, the sensory findings can be counted a pretty good parallel to the neurophysiological records.

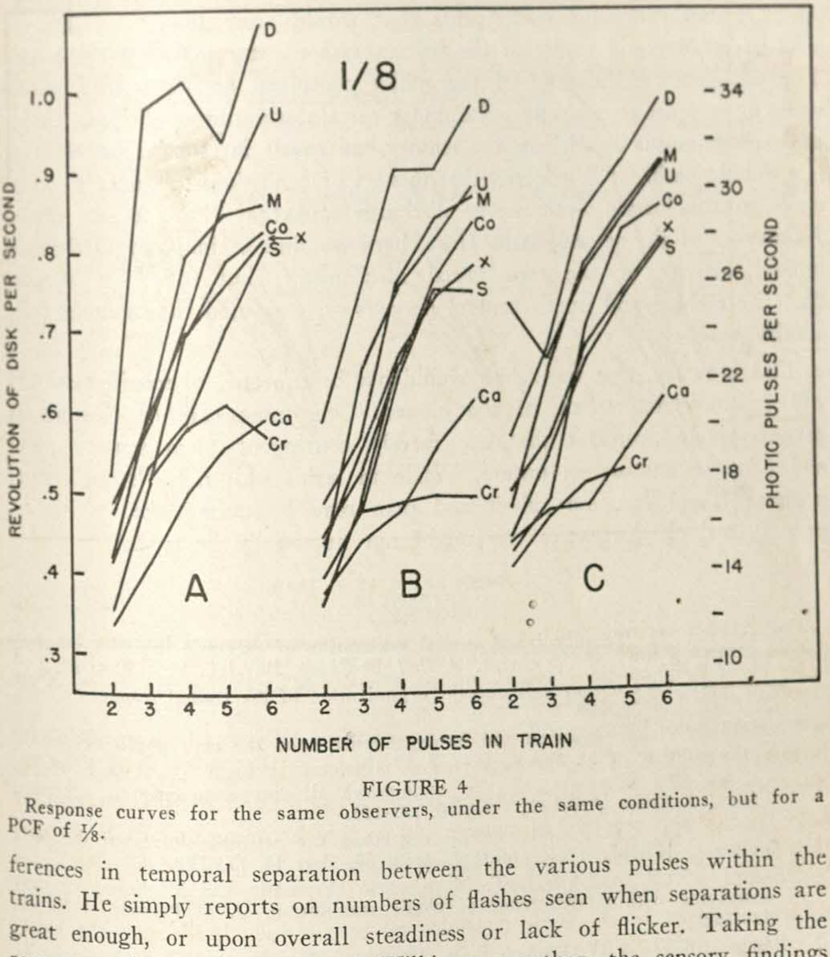


FIGURE 4
Response curves for the same observers, under the same conditions, but for a PCF of ⅛.

The fact that the optic pathway can resolve certain differences between the timing of elements in an extended (repetitive) photic presentation while these are not resolved in full detail in sensory end results is to be taken that the peripheral apparatus of an organism may provide the central mechanism with material which it cannot fully "handle." This was brought out in the findings and implications of certain earlier work (7). There is was shown that the optic nerve discharge of the rabbit to single brief photic pulses contained components that would seem to account for the kind of sensory end result in the human response to similar photic pulses. The optic nerve discharge of the rabbit contained an unexpected second wave in response to a single pulse under certain stimulus conditions. Under the same stimulus conditions, the sensory end result for the human observer is a double flash. It is inferred that this set of conditions produced a second wave in the human optic nerve discharge and that this was the neural forerunner of the second flash. If so, here we find a similarity in the peripheral activity of the two visual mechanisms, but need not suppose that the rabbit's and man's central nervous systems used the input in fully similar ways.

This is to say, that the rabbit would not be expected to manifest sensory end results characteristic of the human. One must carefully distinguish between body processes that are merely contingent upon sensory input and processes sensory in nature. While in terms of purely "sense organ" equipment and function rabbit and man may be quite alike, in higher-order visual characteristics this might not necessarily be true.

REFERENCES

1. BARTLEY, S. H. Temporal and spatial summation of extrinsic impulses with the intrinsic activity of the cortex. *J. Cell. & Comp. Physiol.*, 1936, **8**, 41-62.
2. ———. Some observations on the organization of the retinal response. *Amer. J. Physiol.*, 1937, **120**, 184-189.
3. ———. Visual sensation and its dependence upon the neurophysiology of the optic pathway. *Biol. Symp.*, 1942, **8**, 87-106.
4. ———. The features of the optic-nerve discharge underlying recurrent vision. *J. Exper. Psychol.*, 1942, **30**, 125-135.
5. ———. Some facts and concepts regarding the neurophysiology of the optic pathway. *A.M.A. Arch. Ophthal.*, 1958, **60**, Part II, 775-791.
6. BARTLEY, S. H., & NELSON, T. M. Certain chromatic and brightness changes associated with intermittency of photic stimulation. *J. of Psychol.*, 1960, **50**, 323-332.
7. BARTLEY, S. H., & WILKINSON, F. R. Certain factors in producing complexity of response to a single pulse of light. *J. of Psychol.*, 1953, **35**, 299-306.
8. BASLER, A. Über die Verschmelzung von zwei nacheinander erfolgenden Lichtreizen. *Pflüg. Arch. f. d. Ges. Psychiol.*, 1912, **143**, 245-251.
9. BISHOP, G. H. Cyclic changes in excitability of the optic pathway of the rabbit. *Amer. J. Physiol.*, 1933, **103**, 214-224.

10. Forsyth, D. M., & Chapanis, A. Counting repeated flashes as a function of their number, their rate of presentation and retinal location stimulated. *J. Exper. Psychol.*, 1958, **56**, 385-391.

11. Gebhard, J., Duffy, M., Mowbray, G., & Byham, C. Visual sensitivity to the rate of electrically produced intermittence. *J. Opt. Soc. Amer.*, 1956, **46**, 851-860.

12. Gebhard, J., & Mowbray, G. On discriminating the rate of visual flicker and auditory flicker. *Amer. J. Psychol.*, 1959, **72**, 521-529.

13. Gebhard, J., Mowbray, G., & Byham, C. Difference-limens for photic intermittence. *Quart. J. Exper. Psychol.*, 1955, **7**, 49-55.

14. Gerathewohl, S. Conspicuity of flashing light signals: Effects of variation among frequency, duration, and contrast of the signals. *J. Opt. Soc. Amer.*, 1957, **47**, 27-29.

15. Taubman, R. E. Studies in judged number: II. The judgment of visual number. *J. Gen. Psychol.*, 1950, **43**, 195-219.

16. Wilkinson, F. R. The organization of the visual response. Unpublished doctoral thesis, Michigan State Univer., 1955.

17. ———. The initial perception of pulses of light. *J. of Psychol.*, 1957, **43**, 265-268.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 149-155.

THE GEOCHRONOGRAPH METHOD FOR REPRESENTING NONSTATIONARY ELECTROGRAMS, SUCH AS HEARTBEAT AND RESPIRATION*

Department of Psychiatry, Columbus Psychiatric Institute and Hospital, Ohio State University

Samuel J. M. England and Benjamin Pasamanick

With the imminent advent of procedures for keeping experimental subjects under continuous observation such as radiotelemetering of physiological variables, the pressures to analyze the massive data and to keep the experimenter informed of the natural functioning conditions of the subject indicate that the usual methods for recording and display of the electrograms of repetitive physiological data are defective. Some form of visual representation seems indicated as an integral part of a data analysis system. Figure 1 is representative of a procedure for recording nonstationary repetitive electrograms, such as heartbeat and respiration. The name *Geochronograph* is suggested for this procedure, after the notion of chronogeometry due to A. D. Fokker, cited in Fock (1959) p. 10.

Information obtained from records by the usual methods of recording and rationale for analyzing the data, is acquired at great expenditure of time and effort. Some investigators feel imprisoned by this fact. An improvement in the situation is realized by making use of automatic apparatus for recording and analyzing electrograms, such as the digital computer, relieving the investigator from a struggle with arithmetic detail. The frequency, phase and amplitude, auto and cross correlation functions may be obtained by these means. However these results are not of immediate and compelling utility; they all require a relatively high degree of mathematical sophistication and training for the relationships to be recognized. The analysis of data of the kind that characterizes psychological physiology requires techniques that take cognizance of the fact that the usual notions as to the meaning of time do not apply in the biological system; time does not flow uniformly throughout the various parts of the animal's nervous and physiological activity. This inherent condition imposes upon the experimenter the necessity to resort to special techniques for the analysis of multiple,

---

* Received in the Editorial Office on March 23, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

interacting, nonstationary, time series. The Geochronograph satifies this requirement.

The Geochronograph may be an oscilloscope or paper machine. The two variables, respiration and heartbeat may be superimposed together or heartbeat may be suppressed as is the case in Figure 1. The rate of trace for each line or duration can be continually calculated by the apparatus, interval by interval. In the case of the cathode ray tube, some method of storing

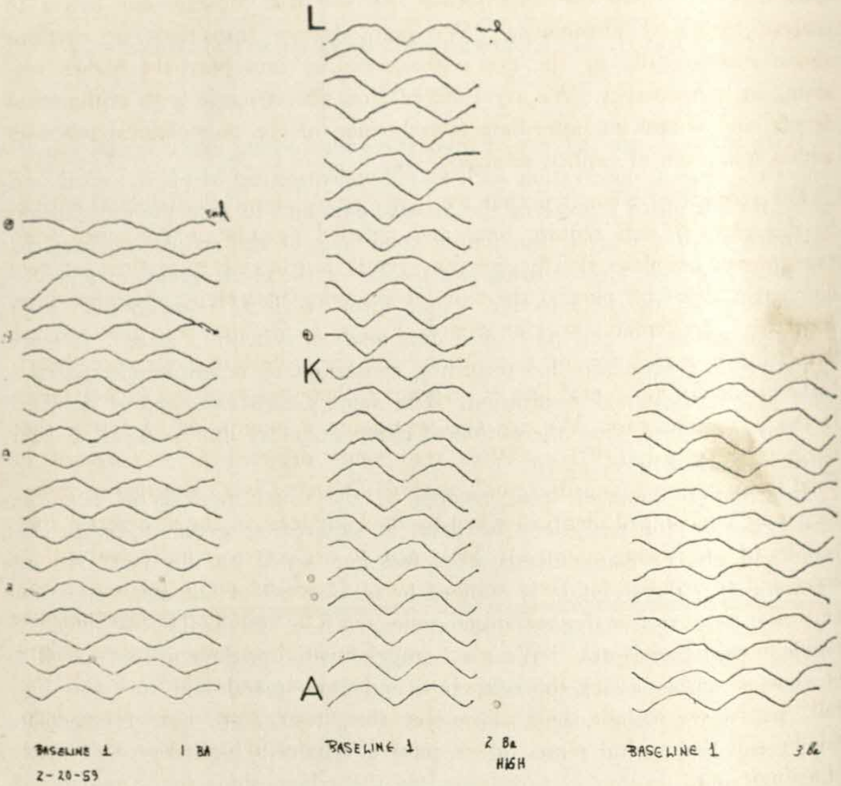


FIGURE 1

Figure 1 shows the electrograph recording of respiration by the method of chromogeometry. Twelve heartbeats, suppressed in the tracing of these records, determine the duration of each line of respiration recording in each of the three experimental sessions. From left to right the recordings were taken at Day 1, Day 14, and Day 28; at Day 14 the subject had been receiving 900 mg/day chlorpromazine since Day 11, and proportionately smaller amounts to zero at Day 1. The subject received no drug after Day 14. The succession of the traces proceed from bottom toward the top.

all previous traces is necessary, until the entire face of the visual display or raster is filled. The picture may then be photographed, erased, and the next duration transcribed. Insofar as this method of recording repetitive electrograms is a successful procedure, it will provide a quick and automatic bridge between the data and its meaning. It will help reorganize notions of space and time with the consequence that we will increase our power to analyze biological phenomena. We may observe transitions in rhythms almost mechanically by the eye without calling into play the higher reasoning of mathematics. We are relieved from the struggle with arithmetical details, and obtain an immediate apprehension of the physiological processes with a minimum of explicit analysis.

The account of what it is that we may observe about physiological entities by this method will require long and detailed elucidation conjoined with experimental results. Briefly we may resort to the interpretations of two different authors for part of the basis of reasoning underlying this procedure. Referring our remarks to the record, Figure 1, we may say that we are witnessing a correlation of respiration with its underlying reference, heartbeat. For every horizontal line of respiration there has recurred 12 heartbeats in this particular case. We are also witnessing a summation of a type that Bergson described (1910). With the events depicted by our record in mind let us try to follow Bergson's analysis: Number is a collection of units; they must be assumed identical when counted (at least in the abstraction that permits of their being counted). We agree to neglect their individual differences and to account for their commonality. If we fix our attention on the particular features of the individual units, we can enumerate them but we cannot obtain their total. "We place ourselves at these two different points of view when we count the soldiers in a battalion and when we call the roll. Either we include them all in the same image, that is we place them side by side in an ideal space, or we repeat 50 times in succession the image of a single one."

In order that the number increase as we advance, we must retain the successive images and set them alongside each of the units which we have pictured for ourselves. It is in space that such a juxtaposition takes place and not in pure duration. Whenever we engage in the activity of repeating 50 times in succession the image of a single one, we shall never have to do more than a single unit, such an occasion constitutes a pure duration and does not lie in space. "Counting material objects means thinking all these objects together, thereby leaving them in space." When adding up the physiologic units in this way, that is, by geochronoscopic presentation, we are not dealing with the durations individually. They will have vanished forever. We are correlating the lasting traces which they seem to have left in space on their passage through it. Our record is constantly on the point of ending and constantly altered by the addition of some new duration. Whereas the inner elements of the durations permeate one another, the durations are exhibited to our regard as distinct as bodies scattered in space; successive durations touch without penetrating.

To describe any phenomenon requires a definite frame of reference or base, a time system for example. If frames, or time systems, are considered one may be inertial or reference. The geometry of this frame of reference is connected with events being measured in the reference space by a system of rules and clocks. At different points along the base we can measure the position of the objects in the space as referred to the base, or rather to the clock at the base. A reference to Figure 1 reveals that our reference frame, the heart time system, is not a *rigid* scaffolding; even though it may be envisaged as having a clock at each of its junctions, all these clocks are not synchronous, some are faster, some are slower, time is a variable magnitude. It is for reasons similar to this and other reasons that the notion of clocks is less useful to psychological physiology than the notion of multiple *time-systems.* An examination of the record of physiological events as represented in Figure 1, makes it apparent that the conceptualizations stemming from A. N. Whitehead (1925) are highly relevant.

In principle a time measurement can be effected with any periodic process. Let us consider the notion of periodicity. Some regularity of recurrence is fundamental, otherwise measurement would be impossible. Beats of the heart recur, breathing recurs. The temporal process can be analyzed into a group of linear serial processes. Each of these linear series is a space-time system. Consider each line or duration in the chart Figure 1. Let one line be an event A, and let B be any part of A, obtained by picking out any one of the events belonging to the series which subdivides A. Note that for the most part A, Figure 1 is subdivided into 3 recurring events of breathing. Recurrence B is then a pattern within the complete pattern A. If B and another recurrence C be two successive slices in the history of duration A, and if C succeeds B, then the pattern C has inherited from pattern B. This evolution is also true of other antecedent relations that form the parts of the life of event C; the other parts, beyond pattern C are inherited from B. "Thus the complete aspects of the antecedent life are inherited as the partial pattern which endures throughout the various periods of the life. The effectiveness of an event beyond itself arises from the aspects of itself which go to form the prehended unities of other events." "It is in this endurance of pattern that time differentiates itself from space." We may here recall Bergson's distinction between counting or summing, and merely enumerating entities. For entities to count up to a sum or collection, there is neglect of their individual differences. The rather general principle enunciated by Fourier, elucidated by Goldman (1948) as well as many others, implies the inheritance from its past condition of every part of a general function or graph, insofar as it may be represented by a collecting or series of recurring durations, each contributing in varying amounts to sum up the whole linear serial process.

When a pattern exhibited in one event is also exhibited in the parts of that event, discriminated according to a certain rule, there is duration, or an enduring object. The rule can be arbitrary, but some rules are trivial as they fail to exemplify a principle or be related to the main stream of abstract ideas that constitute science. In the particular case of the Fourier series, the patterns that are discriminated are periodicities, very often sinusoidal periodicities.

> Apart from the fact of an enduring pattern this principle might be there, but it would be latent and trivial. So to speak, each enduring object discovers in nature, and requires from nature a principle discriminating space from time. Enduring objects are significant of a differentiation of space from time in respect to the patterns ingredient within events. There might be a community (of events) without objects, a bagatelle of transient experience; but there could not be the enduring objects without the community of events with its peculiar patience for them. (Whitehead, 1925).

Now at different stages in its passage an object may vary in the rule for discriminating space from time, in the exhibition of the parts discriminated by a given rule, or both of these changes may be observed. Concerning this point we may consult the recording, Figure 1, and notice that a rule for discriminating the object A, line 1 at the bottom, when compared as if with itself at a later time, such as the duration at line K, results in the conclusion that the object at A has moved to the left with reference to the underlying reference electrogram of heartbeat. "Since objects are nevertheless observed to endure while in a state of motion relative to one another, each endurance is then secured by meanings of space and time not identical from one object to another. Each object may be conceived as at rest in its own space, and in motion throughout any definition of space that does not secure its endurance." The endurance of an object involves the display of a pattern as now realized, and the recurrence of this display again and again and again.

> The duration by the nature of its definition is spatialized; endurance is the repetition of the pattern in successive events; endurance requires a succession of durations. Time is succession of epochal durations. The entities which succeed one another are durations: that which is required for the realization of a pattern. Two objects are mutually at rest when they are utilizing the same rules for discriminating space and time for the purpose of expressing their endurance. If in relative motion their spaces and their times differ (Whitehead, 1925).

One may also observe line L to be of length different from the duration at A, and since the magnitude of extension of each line or duration is reckoned directly from the reference parameter heartrate, then the time system at L differs from the time system at A.

> A body at one stage of its life history that utilizes different meanings for discriminating its space and correlatively different meanings for time, can be said to be in motion relative to itself at another stage. The utilization of different space-time systems means the relative motion of objects. If two particles are relatively at rest, then they are both utilizing the same space-time system. If two objects do not continue either in mutual rest, or mutual uniform relative motion, then at least one of them is changing its intrinsic space-time system (Whitehead, 1925).

The method of observing physiological activity proposed here has many practical advantages: The experimenter can be directly informed of the immediate functioning conditions of the subject under observation; certain types of mathematical analysis are implicitly performed by the manner of presenting the data. The actual quantity of record produced is no more than a fraction of the total usually evolving from electrogram recording while at the same time no information need be lost when compared with the more familiar electrogram recorders. The procedure has the additional advantage of explicitly suggesting its relation to a body of reasoning having the highest abstractive significance. For these combined reasons it seems likely that this form of recording of repetitive electrograms promises to be a useful scientific tool as an adjunct to behavioral experiments.

REFERENCES

1. BERGSON, H. Time and Free Will: An Essay on The Immediate Data of Consciousness. New York: Macmillan, 1910.

2. Fock, V. The Theory of Space Time and Gravitation. New York: Pergamon Press, 1959.
3. Goldman, S. Frequency Analysis, Modulation and Noise. New York: McGraw-Hill, 1948.
4. Whitehead, A. N. Science and the Modern World. New York: Macmillan, 1925.

Columbus Psychiatric Institute and Hospital
University Health Center
473 West 12th Avenue
Columbus 10, Ohio

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 157-160.

CLINICAL PERCEPTION OF SEXUAL DEVIATES*†

Center for Advanced Study in the Behavioral Sciences, Stanford, California

Leo H. Berman and Lawrence Zelic Freedman[1]

In the course of a preliminary survey of antisocial personalities conducted on inmates of a State Prison, clinical impressions reported by the participating psychiatrists were evaluated. We compared their judgments regarding sexual offenders as against those who had committed aggressive and acquisitive acts. Three researchers (two psychiatrists and one psychologist) then rated these randomly sorted descriptions as (1) positive, (2) neutral, and (3) negative. A positive statement was defined as one which indicated general favorable responses such as "robust, vigorous appearance, neatly dressed." A neutral statement was one which expressed no clear value judgment or contained both negative and positive elements such as, "tall, thin, lying in bed, well nourished," or "good-looking, nervous mannerisms, facial tic." A negative statement was one which conveyed disapproval or criticism, extreme examples being "the character of a sex fiend," and "gaunt, weather-beaten, bushy eyebrows and menacing appearance."

The results of the ratings, in which there was consensus by at least two of the three raters is shown in Table 1.[2]

There was a striking difference between the type of descriptive account given by the psychiatrist for the sex- and non-sex-offender. More than half of the sex deviates (52 per cent), were described in negative terms, whereas only one in five (19 per cent), of the non-sex-offenders were so valued. Obversely, two-thirds (67 per cent) of the non-sex-offenders were described by positive statements as compared to slightly less than one-third (30 per cent), of the sex-offender group. Approximately the same percentage of neutral judgments (14 per cent and 18 per cent) were made in each group.

---

* Received in the Editorial Office on April 3, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

† This paper was written with the assistance of the Foundations Fund for Research in Psychiatry.

[1] From the Yale Study Unit of Psychiatry and Law, Yale Medical School.

[2] The reliability of the raters was high. Of 67 items there was complete agreement on 41, and agreement of 2 of the 3 raters on 21 other items with a disagreement never being more than one step. Of the remaining 4 items in which a consensus was not reached, only one contained a two step disagreement.

These differences raised the question as to whether there were countertransference or social-value factors accounting for the differential response.

TABLE 1
RATINGS OF GENERAL STATEMENTS MADE REGARDING PRISONERS

	Sex-Offender		Non-Sex-Offender	
	Number	Percentage	Number	Percentage
Positive	8	30	24	67
Neutral	5	18	5	14
Negative	14	52	7	19
Total	27	100	36	100

$X^2 = 7.50 > .05$ (2df)

To what degree did social and selective perception reflect personal biases unrelated to the clinical reality of the subjects? Alternatively did these findings indicate real differences in the appearance of the sexual offenders as compared to the remainder of the prison population? This was the opinion of Glueck at Sing Sing, who reported that almost 40 per cent of the sexual offenders were "unprepossessing in appearance; and this figure was higher than the control group of the prison population."[3]

We attempted to clarify these questions by setting up a situation in which comparable judgments might be made in the absence of possibly contaminating information concerning the type of offense. We obtained photographs of 200 prisoners, equally divided between sex-offenders and non-sex-offenders, and presented these in a double blind study to 8 psychiatrists. They were asked to distribute these pictures into the four descriptive categories:[4] 1—handsome, 2—average, 3—homely, and 4—ugly or atypical. These photographs were obviously not the equivalent of seeing the man himself, nor were the four descriptive categories identical with the more general clinical statements made by the psychiatrists in our original study. However, this technique permitted control of a number of variables and could be replicated. The photographs were first presented to four psychiatrists who did not know the crime of the prisoner, and several months later they were asked to repeat the experiment, this time with the subjects' crimes listed under each photograph. A second group of psychiatrists first sorted the photographs with the knowledge of the offense, and then after an equal time had elapsed, repeated the sorting with the name of the offense removed from the picture. The large number of subjects, and the time span between the experiments made recognition or perseverating effects unlikely.

---

[3] Glueck, Jr., Bernard C., Final Report-Research Project for the Study and Treatment of Persons Convicted of Crimes Involving Sexual Aberrations, June 1952, June 1955, N.Y. X + 401 pp.

[4] These approximate the descriptive terms used by Glueck.

The results of the ratings by all of the psychiatrists comparing the adjudged appearance of the sexual and non-sexual-offender when the offense was unknown to them is shown in Table 2.

TABLE 2

RATINGS OF PHOTOGRAPHS WHEN OFFENSE IS UNKNOWN

	Sex-Offender		Non-Sex-Offender	
	Number	Percentage	Number	Percentage
Handsome	34	4.3	41	5.2
Average	173	21.3	113	14.3
Homely	491	60.8	593	74.9
Atypical	110	13.6	45	5.7
Total	808	100	792	100

$X^2 = 49.77 > .001$ (df 3)

The figures reveal that even when the offense is not known to the rater, a great number of sexual offenders are rated homely and atypical than are non-sexual offenders, and this difference is statistically significant to the .001 level.

The results of the ratings when the crime was known to the rater is shown in Table 3.

TABLE 3

RATINGS OF PHOTOGRAPHS WHEN OFFENSE IS KNOWN

	Sex-Offender		Non-Sex-Offender	
	Number	Percentage	Number	Percentage
Handsome	22	2.7	34	4.3
Average	500	61.9	602	76.0
Homely	147	18.2	102	12.9
Atypical	139	17.2	54	6.8
Total	808	100	792	100

$X^2 = 110.34 > .001$ (df 3)

The figures show that when the offense is known a greater percentage of sexual offenders are rated in the more ugly and atypical categories, and this difference is statistically significant to the .001 level.

Comparing the distribution of the pictures when the offense is unknown to that when the offense is known reveals a tendency to group the subjects closer to the average. There is also a suggestion of a slight shift of the sex-offenders to the atypical group.

Our results thus indicate that psychiatrists tend to judge sex-offenders as different in appearance from non-sex-offenders. Clinically they may selectively perceive certain aspects of the individual's appearance, bearing, or general relatedness, to which they reacted negatively. However, our photographic study would seem to indicate that there are cues in pictorial representation, to which the psychiatrists react in adjudging sex deviates negatively.

Our observations have not isolated the specific cues in photographs to which the psychiatrists responded. Further insights regarding this might help to clarify some of the problems of aesthetic judgment in relation to interpersonal "tropisms," and may assist in the psychiatrists' attempts to understand better the perceptual factors, both conscious and unconscious which influences his own attitudes in response to patients.

Center for Advanced Study
in the Behavioral Sciences
202 Junipero Serra Boulevard
Stanford, California

Published as a separate and in *The Journal of Psychology*, 1961, 52, 161-177.

THE GROWTH AND EXTINCTION OF EXPECTANCIES IN CHANCE CONTROLLED AND SKILLED TASKS* [1,2]

Department of Psychology, Ohio State University

JULIAN B. ROTTER, SHEPHARD LIVERANT, AND DOUGLAS P. CROWNE

A series of previous studies (Phares, 1957, James & Rotter, 1958, James, 1957) have shown that the growth and extinction of verbalized expectancies for reinforcement are affected by the nature of the task. More specifically, whether or not human *S*s see success in a task as being determined or controlled by chance, random or other factors beyond their control, or see the reinforcement in the situation as an outcome of their own characteristics or skills, appears to have systematic effects on changes in their expectations for future reinforcements. The distinction may be an extremely significant one for learning theory, since it appears that many of the paradigms (such as conditioning, verbal conditioning and probability learning) that are applied to human learning are based partially or entirely on the chance, luck or experimenter controlled situation. However, the behavioral situations to which these models may be applied typically involve mature humans performing in skill tasks.

Differences in learning in two kinds of situations are not solely a matter of degree. The James and Rotter (1958) study indicated a reversal in the rate of extinction of verbalized expectancies under conditions of 50 per cent and 100 per cent reinforcement when the same task was described to the subjects as a skill task, or as a luck task.

Two questions have been raised about these results. One question concerns whether or not the findings are limited to verbalized expectancies and the other deals with the problem of whether or not the findings are in effect the result of the instructions and could not be systematically duplicated by studying the nature of individual tasks. Holden (1960) has recently com-

---

* Received in the Editorial Office on April 5, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was supported in whole by the United States Air Force under Contract No. AF 49(638)-741 monitored by the Air Force Office of Scientific Research of the Air Research and Development Command.

[2] We should like to gratefully acknowledge the assistance of Miss Alice Gordon, Miss Bonnie Strickland and Miss Yvonne Johnson in testing subjects and collating data for this investigation.

pleted a study using betting behavior rather than verbalized expectancies. Extinction was considered to have occurred when *S* was no longer willing to bet on his success. His results replicated the study of James and Rotter in regard to 50 per cent reinforcement and demonstrated that the differences between skill and chance are not limited to measures of verbalized expectancies.

The present study is concerned with a more thorough investigation of the skill-chance dimension as regards both the growth and extinction of expectancies under four different reinforcement schedules. In this study the differences hypothesized in the growth and reduction of verbalized expectancies are based on the nature of the task itself rather than instructions or statements by the experimenter regarding the task.

The studies referred to earlier have taken their point of departure from social learning theory (Rotter, 1954). In this theory, it is assumed that the situation produces or is composed of cues to which the subject has attached expectancies on the basis of previous experience. These are expectancies that specific behaviors will lead to specific consequences or reinforcements. The expectancies themselves are determined by the past history of reinforcement with these same cues, and generalized expectancies which may relate either to the nature of the behaviors involved, the reinforcements, the nature of the cues, or all three. Generalized expectancies for specific kinds of behaviors leading to similar reinforcements are described by the concept of freedom of movement. Some generalized expectancies, however, cut across specific kinds of reinforcements and have to deal with what might broadly be called problem solving skills or perhaps what Harlow (1949) has referred to as higher level learning skills. Such generalized expectancies might include, for example expectancies related to concepts of causality, delay of gratification, the consequences of trial and error behavior, looking for alternatives, etc. There is some evidence to support the notion (Rotter, Seeman & Liverant, 1961) that such a generalized expectancy is present in adult subjects in regard to the implications for the future occurrence of reinforcements which appear to result as a function of their own stable characteristics versus chance.

It should follow from the above assumptions that in a given task if *S* regards the task as being more controlled by luck or chance he is likely to "learn" less about future behavior-reinforcement sequences from the occurrence of the reinforcement following the behavior. In other words, his expectancies for future reinforcement are likely to change less when he regards the occurrence of the reinforcement to be beyond his control. It was hypothesized, then, that if a subject regards success in a particular task as determined by luck, chance or external control his expectancies for future positive reinforcement would, (1) rise less after positive reinforcement, (2) fall less after negative reinforcement, (3) be more subject to the gambler's fallacy (the occurrence of a negative reinforcement increases the probability of a subsequent positive reinforcement or vice versa) and (4) in an extinction series (continuous negative reinforcement) would be slower to extinguish as a consequence of (2) and (3). In the case of 100 per cent or near 100 per cent reinforcement the individual who regards the reinforcement as dependent upon some kind of external control would assume early in the extinction trials either that luck had turned against him or that the experimenter was no longer rewarding the same behavior.

On the other hand, a subject who regarded reinforcement as a function of his own skill would extinguish more slowly under 100 per cent reinforcement than under partial reinforcement because it would take him longer to accept the fact that his presumably stable skills had diminished. He is not anticipating external changes in the situation which will account for failure. The present study is designed to provide one or more tests of all these hypotheses.

A. Selection of Tasks

In order to test these hypotheses, we needed to find one task that *S*s in a given culture would regard as being primarily luck or chance determined, and one in which the subject would feel that his performance was primarily a matter of skill. The two tasks at least had to be discriminable on a continuum of skill versus chance control. The design produced some difficult requirements. In order to make the two tasks comparable the same series of reinforcements during an acquisition or growth of expectancy phase would have to be presented. Consequently, both tasks must be controlled by *E*, but in one case *S* should believe that performance is a matter of his own skill and in the other that it is a matter of luck. In the extinction phase it would be necessary for *S*, under one condition, to change from 100 per cent success trials to no success without being suspicious that *E* was, in fact, controlling the task.

A number of tasks were tried out in an attempet to meet these conditions and finally tasks were found for female *S*s which appeared to work with a minimum of suspicion of experimenter control.

The task selected as a skill controlled or internally controlled task was adapted from Cromwell's (1959) modification of Sky's (1950) apparatus. We will refer to it here as the Vertical Aspiration Board. The apparatus is shown in Figure 1. In this task the subject is asked to pull on a string connected with a platform on which a steel-bearing ½" in diameter is placed. The subject is required to lift the platform smoothly to a specified height before the ball rolls off. The ball itself is held on the platform by an electromagnet which can be switched on and off by a silent knee switch underneath the table. Since the platform itself slopes forward, the ball falls off instantaneously after the current is cut off by the knee switch. Complete control of scores is thus obtainable. The apparatus is constructed so that there are no wires in evidence, the current being supplied to the moving platform via contact with a strip of copper not visible to the subject. The current is supplied by a dry cell hidden in a false bottom on the table. A detailed description of the apparatus may be obtained from the authors.

The chance or luck task is similar to the one used by James & Rotter (1958) in their previous report. The subject is shown a number of stacks of five cards with X's and O's on them. His problem is to guess each card in a group and to obtain a better than chance (four or five) number of correct guesses in order to succeed. The cards are placed in a simple tachistoscope and immediately following each guess a window raises and the subject can see whether he is correct or not. However, none of the actual stacks of cards are used by the experimenter hidden behind the tachistoscope. Instead, a single card with an X and an O on it is moved back and forth silently before the window rises and immediately after the subject has stated his guess. With a little practice our experimenters carried this off so well that in our large group of subjects not a single one verbalized any suspicion that the experimenter was controlling his success in guessing. While it is true in this task that some subjects in some degree may have felt that a true ESP skill might be involved, our over-all assumption was that such a task for the group as a whole would be considered to be much more determined by chance, luck, or random factors than the task utilized as a skill task.

B. SUBJECTS

In pre-testing the skill tasks a number of the men expressed some suspiciousness, primarily in the 100 per cent reinforcement group. Consequently, it was decided to run this experiment entirely with female *S*s. The men who did express suspiciousness did not know what was happening but felt that in some way the experimenter was controlling the outcome. The design called for eight groups or conditions of twenty *S*s each. All *S*s were from the

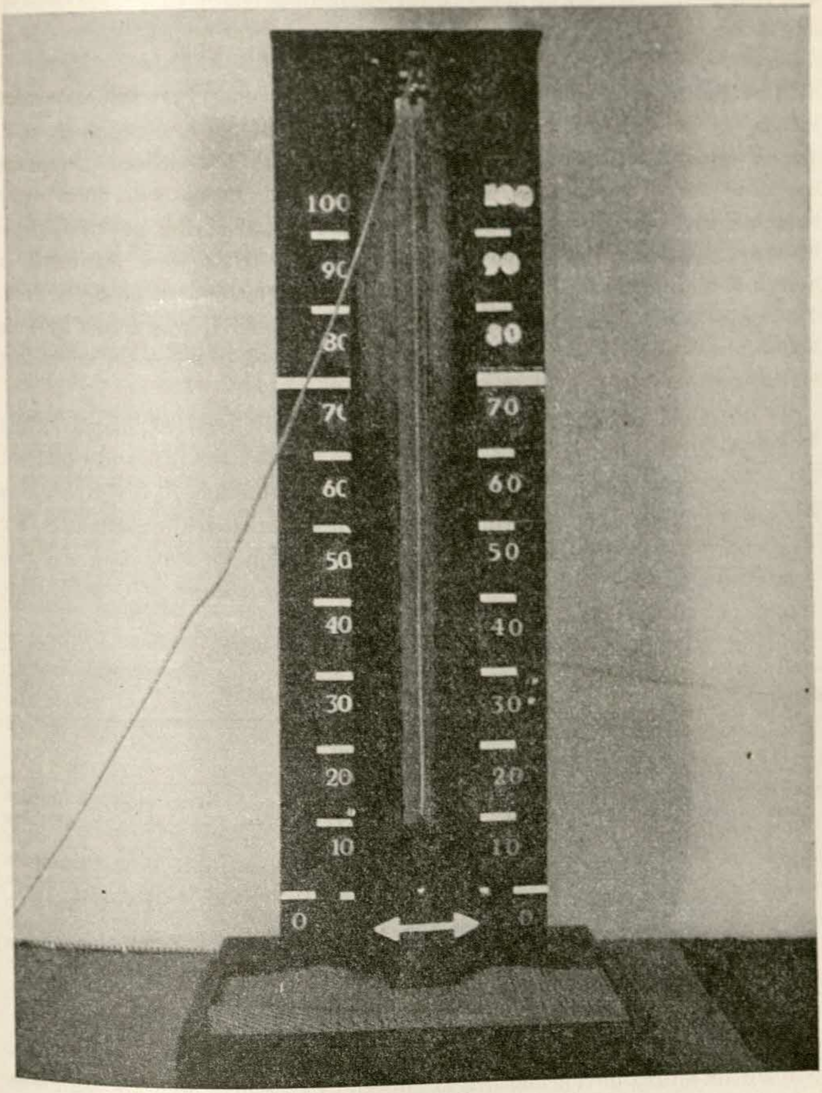


FIGURE I
THE VERTICAL ASPIRATION BOARD

elementary psychology courses at The Ohio State University where *S*s are required to serve in one or more experiments although they have a choice of the type of the experiment in which they wish to participate.

C. Design and Procedure

The 160 *S*s were randomly assigned to eight groups. For each task there was a 100 per cent, 75 per cent, 50 per cent, and 25 per cent positively reinforced group. Positive reinforcement and failure are defined herein as reaching or not reaching a criterion score. Each group was given eight training trials and in each group the first and last trial was positively reinforced. Following the training trials, extinction trials were instituted in which *S* continuously failed to achieve the necessary standard on every trial. In summary the experimental design, shown in Table 1, was a two by four factorial for which analysis of variance was employed to assess the dependent variables.

If the *S* did not extinguish earlier the experiment was terminated on the 50th extinction trial. In every case, after being shown the task and explained what was required, each *S* stated on a scale of 0 through 10 his expectation of succeeding on the following trials. These verbalized expectancies for success provide the data for this study. The criterion for extinction was two consecutive trials with a verbalized expectancy for success of 0 or 1. Instructions and the detailed procedure are given below.

TABLE 1
Experimental Design

			Training Trials								
Group		N	1	2	3	4	5	6	7	8	
25%	Chance	20	+	—	—	—	—	—	—	+	Extinction
25%	Skill	20	+	—	—	—	—	—	—	+	Extinction
50%	Chance	20	+	—	—	+	+	—	—	+	Extinction
50%	Skill	20	+	—	—	+	+	—	—	+	Extinction
75%	Chance	20	+	+	—	+	+	+	—	+	Extinction
75%	Skill	20	+	+	—	+	+	+	—	+	Extinction
100%	Chance	20	+	+	+	+	+	+	+	+	Extinction
100%	Skill	20	+	+	+	+	+	+	+	+	Extinction

1. *Vertical Aspiration Board Instructions*

This experiment is to see how well you can succeed in raising this ball without its falling off (demonstrate) and also to see how accurate you are in estimating your success.

The object of this task is for you to try by pulling this string to raise the ball on the platform as high as possible before the ball drops off. You will be given a number of trials. The apparatus is built with a very slight tilt forward so that the ball is more likely to fall off the platform the higher it is raised. Of course if you raise the platform very quickly the ball can't drop off because of the momentum, therefore the platform must be raised slowly. Now, in order to be clearly successful you must score 80 or better on a given trial.

Before each trial I would also like you to estimate how certain you are that you can succeed, i.e., get 80 or better. You are to estimate your degree of certainty of success on a scale going from 0 to 10. For example, if you feel fairly sure that you will succeed, that is, that you will get 80 or higher, you may rate yourself with a nine or ten. If you feel moderately sure that you will succeed you may rate yourself with a 4, 5, or 6. If you feel pretty sure you will *not* be successful, i.e., not get 80, you may rate yourself with a 0 or 1. Use any numbers on the scale (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10) to indicate how successful or unsuccessful you think you will be. It is important that you select your estimates carefully on the 0 to 10 scale and that they correspond closely with how certain you *really* are. They should be an accurate description of the degree to which you really feel that you will or will not succeed.

Do you have any questions so far? All right ——, do as well as you can. Now, before we begin, make an estimate on the scale from 0 to 10 as to what you think the likelihood is of your getting 80 or better on the first trial.

2. *T scope Instructions*

This experiment is to see how well you can do at telling me beforehand which of two kinds of cards will be exposed in this opening (pointing) and also to see how accurate you are in estimating your success.

In this box (show box) we have a large number of cards marked with either an X or an O (show one of each). These cards are divided into sets of five. Each set has been *shuffled* and there are not necessarily the same number of X's and O's in each set. Each time I will select *at random* one of these sets of five cards and place it behind this closed shutter (point). (Remove box.) You are to tell me whether the top card is an X or an O. After you tell me, this shutter will open and you will know whether you were right or wrong. I will remove that card and the next card in the set will be in front of the closed shutter and again you will tell me whether it is an X or an O. In this way we will go through a number of these sets of 5 cards. I will also be keeping score and will let you know how well you did at the end of each trial, that is, at the end of a set of five cards.

Now in order to be successful on a trial you must get at least four or more cards right out of five. Four or five cards correct will mean that you have succeeded. Any number of cards correct below four will mean that you have *not* succeeded.

The remainder of the instructions are identical with those of the Vertical Aspiration Board.

D. RESULTS

1. *Training*

Our over-all hypothesis specified that when *S* perceives a task to be controlled by external forces he "learns" less from the occurrence of a reinforcement than if he perceives the reinforcement in a situation to be exclusively, or to a greater degree, a function of his own behavior and efforts. Four tests of this hypothesis were made on the training trial data.

The first test applied was the change in verbalized expectancies from the first to the second trial. All eight groups were positively reinforced on the first trial. At the beginning of the second trial, then, all the subjects were comparable in percentage of reinforcement and it was possible to compare the four skill task groups with the four task groups in the degree to which a positive reinforcement resulted in a rise in expectancy for the subsequent trial.

A second test of our over-all hypothesis was made by comparing the expectancies at the end of the training sequence. In this case, four separate comparisons were made between the 25 per cent, 50 per cent, 75 per cent, and 100 per cent skill and chance groups. If changes in expectancies were differential for the two tasks, cumulative effects of the eight training trials could best be shown at the beginning of the 9th trial.

The third test of differences in approach to learning in the two kinds of situations was a measure of unusual shifts, that is, shifts up after failure or down after success in the verbalized expectancy for future success. It was hypothesized that such shifts would be greater in a situation perceived as being controlled by external factors or chance.

A fourth test involved a more complicated measure. In this case we were interested in the extent to which the individual varied his expectancies as a function of his new experience in a manner consistent with the previous reinforcement. The measure which we will refer to as Total Amount of Change was then taken simply to include all of the appropriate shifts in expectancy which followed from the previous reinforcement. That is, the total amount or sum of relevant shifts or changes up after success and down after failure in verbalized expectancies was computed for each subject. In this measure, unusual shifts (see above) were excluded. These four tests of differences between skill and chance groups are obviously not independent.

The means and standard deviations for expectancy statements during the training trials are given in Table 2. In order to make comparisons between groups for subsequent trials it is first necessary to demonstrate that there are no significant differences in initial means which might account for later differences. Consequently, an analysis of variance was made for the initial statements of expectancy for the eight groups. The over-all *F* was not significant for initial expectancies. As can be very clearly seen from Table 2, the amount of rise for *S*s in the skill tasks for all four groups was considerably greater following the first positive reinforcement than for any of the chance groups. Here the analysis of variance was highly significant (F = 69.91; df = 1 & 152; *p* = < .001). The consistency of the differences between

TABLE 2
MEANS AND STANDARD DEVIATIONS OF EXPECTANCY STATEMENTS DURING THE TRAINING TRIALS

Group			Training Trials								
			1	2	3	4	5	6	7	8	9
25%	Chance	M	4.0	5.4	4.5	3.3	3.3	3.1	4.0	2.9	4.5
		SD	2.0	2.1	2.0	1.6	1.5	1.8	1.7	2.1	2.0
25%	Skill	M	4.0	7.8	7.2	6.0	4.6	4.3	4.4	3.5	5.8
		SD	2.1	1.7	1.6	1.9	2.4	2.6	2.2	2.1	2.0
50%	Chance	M	3.9	5.1	4.5	4.6	5.6	6.2	5.9	5.1	6.0
		SD	1.2	1.6	1.2	1.3	1.7	1.6	1.7	1.6	1.9
50%	Skill	M	3.5	7.3	6.7	5.3	7.6	7.5	7.1	6.3	7.4
		SD	1.6	1.6	1.4	1.7	1.3	1.6	1.4	1.5	1.2
75%	Chance	M	3.9	5.6	7.0	5.8	7.6	8.1	8.8	6.7	7.0
		SD	2.0	1.3	1.2	1.5	1.0	1.2	.9	1.4	1.5
75%	Skill	M	3.2	7.4	7.9	7.0	8.2	8.3	8.4	7.6	8.2
		SD	1.8	1.9	1.6	2.0	1.7	1.5	1.5	2.1	1.7
100%	Chance	M	3.8	5.6	6.7	7.5	8.4	8.7	8.8	9.0	9.0
		SD	2.0	1.8	1.8	2.0	1.6	1.7	1.5	1.4	1.5
100%	Skill	M	3.8	7.8	8.4	8.4	9.3	9.3	9.1	9.2	9.4
		SD	1.7	1.1	.9	1.0	.7	.7	.9	.8	.8

comparable groups is striking. It is clear that when faced with the tasks which most of them regarded as one involving skill to begin with, the first reinforcement led to a much sharper rise in expectancy for future success.

The analysis of variance for expectancies at the end of the training trials yielded a significant *F* for the difference between skill and chance groups and for reinforcement schedules. (Skill-Chance, F = 17.11; df 1 & 152; $p = < .001$.) (Reinforcement schedules, F = 42.96; df = 3 & 152; $p = < .001$.) In every case, higher expectancies were stated at the end of the training sequence for the skill groups than for the chance groups. These differences are significant for the 25 per cent ($p = < .05$), 50 per cent ($p = < .01$), and 75 per cent ($p = < .05$) reinforcement schedules but not for the 100 per cent groups where both groups were beginning to asymptote at the upper limit of the expectancy scales.

The number of unusual shifts in the eight groups were not subjected to analysis of variance and the hypothesis of difference was rejected on inspection. There were a total of 47 unusual shifts in the four skill groups and 46 in the four chance groups. Only 42 per cent of the subjects had one or more unusual shifts.

An analysis of variance was computed on the total amount of change in the direction of the previous reinforcement which yielded, as shown in Table 3, a significant interaction.

TABLE 3
ANALYSIS OF VARIANCE OF TOTAL CHANGE IN THE DIRECTION OF PREVIOUS REINFORCEMENT

Source	df	MS	F
Skill vs. Chance	1	117.30	6.69*
Reinforcement Schedules	3	114.62	6.54**
Skill vs. Chance × Reinforcement Schedules	3	78.36	4.47**
Within (error)	152	17.53	
Total	159		

* $p < .02$
** $p < .01$

Table 4 presents the means and standard deviations for the eight groups and the significance of the difference for the skill versus chance paradigms. Again it can be seen that there was more relevant changing of expectancies in all four skill groups than in the comparable chance-task groups. These differences are significant for the 25 per cent and 50 per cent groups, but not for the 75 per cent and 100 per cent groups where there was, in general, less shifting since more subjects reached an asymptote for their verbalized expectancies before the end of the training series.

TABLE 4
MEANS, STANDARD DEVIATIONS, AND DIFFERENCES BETWEEN AND WITHIN SKILL AND CHANCE FOR THE AMOUNT OF CHANGE IN THE DIRECTION OF PREVIOUS REINFORCEMENT

Group	N	Mean	SD	Difference	*t*
Between Skill and Chance					
25% Skill	20	11.95	4.55	4.55	3.44***
25% Chance	20	7.40	4.11		
50% Skill	20	11.10	4.37	3.35	2.53**
50% Chance	20	7.75	4.78		
75% Skill	20	9.10	5.44	1.70	1.28
75% Chance	20	10.80	4.41		
100% Skill	20	6.65	2.49	0.65	0.49
100% Chance	20	6.00	2.37		
Within Skill and Chance					
25%- 75% Skill				2.85	2.15**
50%- 75% Skill				2.00	1.51
50%-100% Skill				4.45	3.36***
75%-100% Skill				2.45	1.85*
75%- 25% Chance				3.40	2.57**
75%- 50% Chance				3.05	2.30**
75%-100% Chance				4.80	3.63***
50%-100% Chance				1.75	1.32

* $p < .10$
** $p < .05$
*** $p < .01$

2. *Extinction*

The major measure of the differences in learning in the two kinds of situations was the number of trials to extinction under continuous failure to achieve the specified level of performance. Figure 2 shows the mean number of trials to the extinction criterion of 0 or 1 expectancy for two continuous trials for the four skill and four chance groups. From Figure 2 and the analysis of variance given in Table 5 it can be seen that a complete reversal takes place in trials to extinction for skill versus chance groups. Neither the skill-chance dimension itself or the reinforcement schedule is significant alone, but the interaction between reinforcement schedule and the skill chance variable is highly significant. Homogeneity of variance was established by the F Max test (Walker & Lev, 1953).

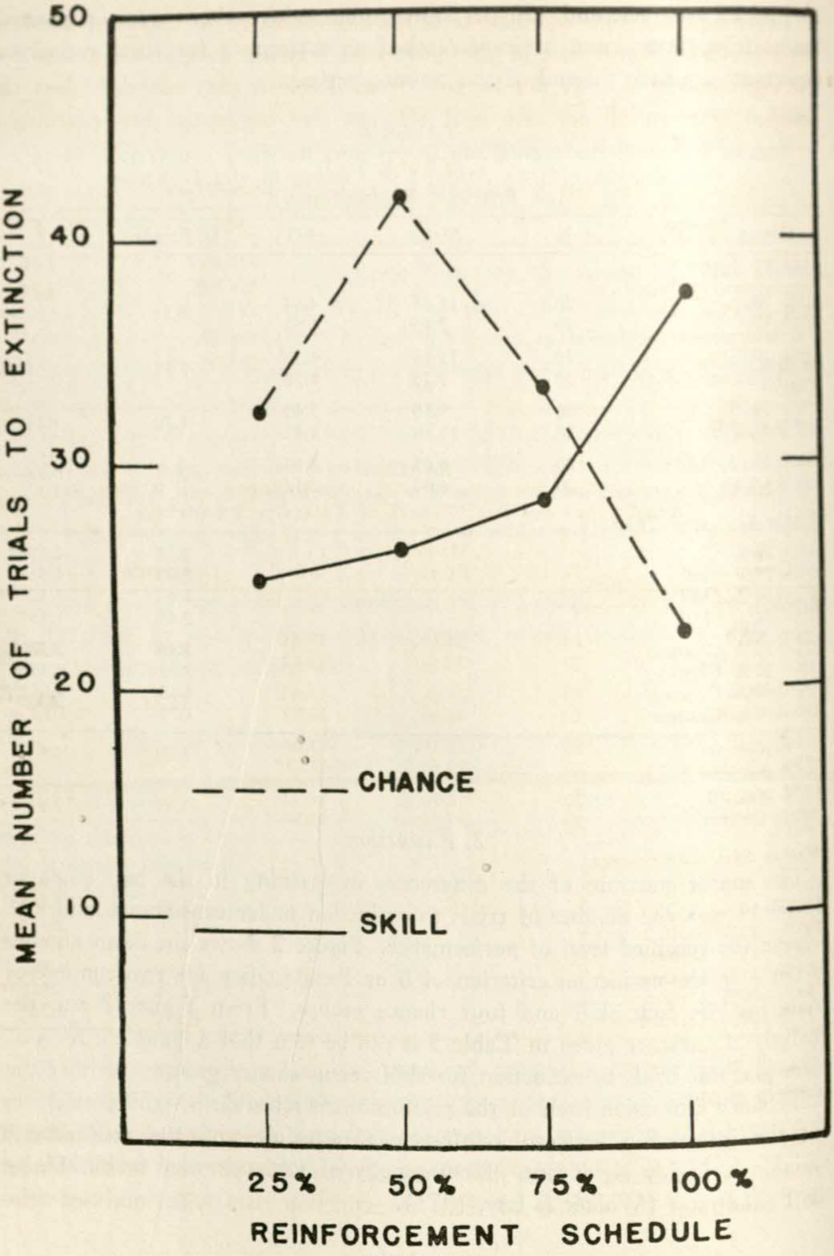


FIGURE II
Mean Number of Trials to Extinction for Skill and Chance Groups Under the Four Reinforcement Schedules

Table 6 presents the significance of the differences of the means of standard deviations and significance of the differences between comparable groups for the eight conditions. It can be seen from Table 6 that not only does the reversal between 50 per cent and 100 per cent or partial and continuous reinforcement take place but the differences between comparable chance and skill groups are significant at both the 50 per cent and 100 per cent levels. Under skill conditions the number of trials to extinction rises with increased rein-

TABLE 5
ANALYSIS OF VARIANCE OF TRIALS TO EXTINCTION

Source	df	MS	F
Skill vs. Chance	1	409.60	1.41
Reinforcement Schedules	3	214.30	0.74
Skill vs. Chance × Reinforcement Schedules	3	1835.03	6.30*
Within (error)	152	291.06	
Total	159		

* $p < .001$

TABLE 6
MEANS, STANDARD DEVIATIONS, AND DIFFERENCES BETWEEN AND WITHIN SKILL AND CHANCE FOR THE NUMBER OF TRIALS TO EXTINCTION

Group	N	Mean Trials	SD	Difference	*t*
Between Skill and Chance					
25% Skill	20	24.60	19.64	8.00	1.48
25% Chance	20	32.60	19.12		
50% Skill	20	26.30	13.65	15.75	2.92***
50% Chance	20	42.05	16.97		
75% Skill	20	28.10	17.06	5.00	0.93
75% Chance	20	33.10	17.71		
100% Skill	20	38.70	16.00	15.95	2.95***
100% Chance	20	22.75	15.56		
Within Skill and Chance					
100%- 75% Skill				10.60	1.96*
100%- 50% Skill				12.40	2.30**
100%- 25% Skill				14.10	2.61**
50%- 25% Chance				9.45	1.75*
50%- 75% Chance				8.95	1.66
50%-100% Chance				19.30	3.58***
75%-100% Chance				10.35	1.92*

* $p < .10$
** $p < .05$
*** $p < .01$

forcement in the training sequence. Under chance conditions, extinction is protracted at the 50 per cent level and, relative to all other groups, quite short at the 100 per cent level.

E. Discussion

Although the number of training trials was relatively short, major differences appeared between the groups tested on the skill apparatus and on the chance apparatus. These differences must be attributed to the ways the groups perceived the two tasks as a function of common cultural experience. The considerably greater rise for the four skill groups following the first positive reinforcement makes sense in line with our original hypothesis that the subject will change his expectancies for future reinforcement to a larger extent if he attributes the occurrence of this reinforcement to some permanent characteristic of himself, such as a skill.

The second test involving expectancies at the end of the training sequence likewise showed differences in the same direction between all of the comparable reinforcement schedule groups. It is clear enough why this should be true of the 75 per cent and 100 per cent groups, but not equally clear of why it would be true of the 25 per cent and 50 per cent groups. The 25 per cent group in particular, under skill conditions, should not have been expected to rise to an expectancy level so much higher than the actual frequency of previous reinforcement. It seems likely that the best explanation for this rise is that in all the skill groups with the exception of the 100 per cent group (in which some people typically asymptote at an expectancy level of nine presumably not being willing to state the highest possible expectancy), that an overestimation took place as a result of previous experience with skill tasks. A level of aspiration study by Sutcliffe (1956) in which he compares skill and guessing random events gives strong support to this interpretation. In other words, if the subject perceived the task as one of skill, he habitually expected to do better as a function of practice on subsequent trials. Having succeeded once he anticipates he can do it again because he possesses the necessary skill. Although there may be other explanations for these findings, this appears to be the most logical alternative.

The failure to find differences in unusual shifts in the chance versus the skill groups cannot readily be explained. Phares (1957), using 13 trials and a single task, in which instructions differentiated his chance from his skill groups, was able to find a difference approaching significance ($p = .07$). It appears that for the tendency to commit the gamblers fallacy, at least in regards to this measure of unusual shifts, verbally designating a task as being controlled by skill or by luck appears to be a more effective method of producing differences than taking two tasks differentiated on the basis of presumed common cultural experiences.

The fourth test of total changes in verbalized expectancies in the two comparable groups yielded results as predicted. What this difference between skill and chance groups indicates is that under skill conditions subjects lower their expectancies more after failure and raise them more after success over all the eight trials. In other words, past reinforcement is a clearer cue to future reinforcement when the subject believes that his own skill determines outcome.

One of the more interesting aspects of the extinction data is that differences between chance and skill groups is much smaller at 25 per cent and 75 per cent reinforcement schedules than at 50 per cent and 100 per cent. One hypothesis to account for this finding is that the 25 per cent and 75 per cent group in the chance task were being rewarded or reinforced less or more often than could be accounted for by chance alone. The frequency of reinforcement itself may tend to make the task appear more like a skill task. In a recent dissertation, Blackman (1960) was able to demonstrate that non-obvious patterning of reinforcements apparently affected the perception of a task as skill or chance controlled. The greatest difference among the partially reinforced groups is at the 50 per cent level where an even distribution of successes and failures in the chance task might enhance the notion that the task is in fact being controlled by luck. This does not account, of course, for the fact that the 100 per cent chance group would not also see the task as skill controlled because of the high frequency of reinforcement. Our hypothesis here is that the nature of the task is such that although the subjects may tend to eventually regard the task as skill determined during the training trials, the non-occurrence of the reinforcement in the first few trials of extinction quickly strengthens the original notion that it is a luck task or one subject to external control. That is, they quickly hypothesize that the experimenter has, in fact, changed something about the task or that the early successes were in fact a matter of luck which has run out. Such an explanation invokes two different principles to account for the differences between chance and skill groups. Although this explanation is not parsimonious it has at least the virtue of fitting the data.

F. Summary and Conclusions

Previous studies by Phares (1957), James and Rotter (1958), and James (1957), have demonstrated that major differences in the growth and extinction of verbal expectancies take place when a task is perceived as one being controlled by chance or luck versus being controlled by skill or internal factors. That is, these differences obtain when *S* anticipates that reinforcement occurs as a function of something external to himself in contrast to his belief that reinforcement is a function of his own behavior. Obviously, tasks could differ along this dimension in degree as well as kind. These previous studies suggest that the differences in behavior in the two kinds of tasks are not merely a matter of degree but, in fact, different principles of learning apply. This is shown particularly in the reversal of partial versus 100 per cent reinforcement conditions in the chance versus skill groups, with the typical finding (Lewis, 1960) of superiority of resistance to extinction in the partially reinforced condition found only in the externally controlled or chance task. Such findings are of considerable importance both in learning theory and in the application of learning principles to a complex life situation.

However, the findings described above were all based on the use of a single task or tasks in which the difference between chance control and internal control was established by differential instructions. In the present study an attempt was made to study the differences in the growth and extinction of expectancies using two different tasks which we hypothesized would be regarded as skill and chance controlled tasks on the basis of the previous cultural experience of the subjects, rather than by instructions. Eight groups of subjects were studied with an apparatus involving a motor skill of steadiness and one of guessing cards. For each task a group of 20 subjects were given 25 per cent positive reinforcement, 50 per cent positive reinforcement, 75 per cent positive reinforcement, and 100 per cent positive reinforcement, for eight training trials. On nonreinforced trials the subject failed to achieve a criterion score. Following this all groups were given continuous failure until they extinguished by stating expectancies for future success of 1 or 0 for two consecutive trials.

The general hypothesis tested was that under chance conditions the occurrence of reinforcement leads to less learning in that previous experience is a less stable predicter of the future occurrence of the event. In general, the results strongly supported the two hypotheses. First, that under skill conditions positive and negative reinforcement leads to greater increments and decrements in verbalized expectancies. Second, that the extinction of expectancies under continuous negative reinforcement will reverse under chance and skill conditions so that 50 per cent is more resistant to extinction than 100 per cent reinforcement under chance conditions, and 100 per cent reinforcement is more resistant to extinction than 50 per cent reinforcement under skill conditions.

REFERENCES

1. BLACKMAN, S. Some factors affecting perception of events as chance determined. Unpublished doctoral dissertation. The Ohio State University, 1960.
2. CROMWELL, R. A methodological approach to personality research in mental retardation. *Amer. J. Ment. Defic.*, 1959, **64**, 333-340.
3. HARLOW, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51-65.
4. HOLDEN, K. B. A non-verbal measure of differences in extinction in skill and chance interactions. Unpublished doctoral dissertation, The Ohio State University, 1960.
5. JAMES, W. H. Internal versus external control of reinforcement as a basic variable in learning theory. Unpublished doctoral dissertation. The Ohio State University, 1957.
6. JAMES, W. H., & ROTTER, J. B. Partial and one hundred per cent reinforcement under chance and skill conditions. *J. Exp. Psychol.*, 1958, **55**, 397-403.
7. LEWIS, D. J. Partial Reinforcement: A selective review of the literature since 1950. *Psychol. Bull.*, 1960, **57**, 1-28.
8. PHARES, E. J. Expectancy changes in skill and chance situations. *J. Abnorm. Soc. Psychol.*, 1957, **54**, 339-342.
9. ROTTER, J. B. Social Learning and Clinical Psychology. New York: Prentice-Hall, 1954.
10. ROTTER, J. B., SEEMAN, M. R., & LIVERANT, S. Internal versus external control of reinforcement: A major variable in behavior theory. In (N. F. Washburne, Ed.) Decisions, Values and Groups, Vol. 2., London: Pergamon Press, 1961.
11. SKY, A. W. An apparatus for a frustration task. *Australian J. Psychol.*, 1950, **2**, 116-120.
12. SUTCLIFFE, J. P. Random events as a function of belief in control. *Australian J. Psychol.*, 1956, **8**, 128-139.
13. WALKER, H. M., & LEV, J. Statistical Inference. New York: Henry Holt, 1953.

Department of Psychology
Ohio State University
Columbus 10, Ohio

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 179-183.

THE RELATIONSHIP BETWEEN RATIONAL AND IRRATIONAL AUTHORITARIANISM*[1]

School of Psychology, Georgia Institute of Technology

STANLEY A. RUDIN

In one of the earliest of contemporary theoretical discussions of the social psychological nature of authority, Fromm (1947) distinguished between *rational* and *irrational* authority.

> *Rational authority* has its source in competence. The person whose authority is respected functions competently in the task with which he is entrusted. . . . (It) not only permits but requires constant scrutiny and criticism . . . (p. 9).
>
> . . . Humanistic ethics are not incompatible with rational authority (p. 10).
>
> The source of *irrational authority* . . . is always power over people. This power can be physical or mental . . . . Criticism . . . is forbidden . . . . (p. 9).
>
> Rational authority is based upon the equality of both authority and subject, which differ only with respect to the degree of knowledge or skill in a particular field. Irrational authority is by its very nature based upon inequality, implying difference in value (pp. 9, 10).

Elsewhere (Fromm, 1941), rational authority is likened to the teacher-student relationship, and irrational authority to that between master and slave.

Probably because of the immediate danger from groups and philosophies exemplifying irrational authority at that time, research since Fromm's early work has centered on irrational authority and ignored rational authority. The builders of the *F*-scale (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950), although acknowledging that the thinking of Fromm, among others, had influenced their own (footnote 1, p. 231), concentrated on defining and measuring the tendency to accept irrational authority, ignoring its complement.

---

* Received in the Editorial Office on April 5, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The data described herein were collected as part of a larger study performed at the Personnel Laboratory, Det. #1 of Wright Air Development Division, Lackland Air Force Base, Texas. All conclusions and opinions are those of the author, and do not necessarily represent those of the U.S. Air Force.

Rudin (1958), has called attention to the problem of academic anti-intellectualism, one of whose symptoms is a failure by the student to perceive the teacher as a rational authority worthy of respect and obedience. Teachers and youth workers in recent years also have begun to express concern over the anti-authority sentiments and hostility toward *all* authority of increasing numbers of young people. Correlated with this tendency is another to exhibit exaggerated conformity to the norms of their peer group. It would seem desirable, therefore, to begin the systematic study of the *tendency to accept rational authority*. This study asks two basic questions: (a) can acceptance of rational authority be operationally defined and measured as a unitary trait? and (b) does it correlate with the tendency to accept irrational authority?

The first aims at finding whether empirical research using a simple pencil-paper questionnaire is possible. The second is prompted by the assumption frequently made (implicitly or explicitly) by youth workers, that training youths to resist *all* forms of authority might make them a bit wild, but will at least prevent their acceptance of irrational authority. This assumption predicts a significant negative relationship between the two attitudes.

A. METHOD

1. *Instruments*

The California F-scale was assumed to measure the tendency to accept irrational authority since its authors acknowledge Fromm's influence and their description of the authoritarian's personality conforms closely to that of Fromm's irrational authority. The items used were from Table 7 (VII), Forms 45 and 40: 1, 12, 37, 41, 4, 8, 21, 23, 42, 44, 13, 19, 25, 27, 34, 39, 9, 31, 16, 26, 29, 33, 2, 6, 43, 18, 35, 38. To give 30 items, two more were added from earlier versions of the test: item 41, from Table 4 (VII), Form 60; and item 9, page 238, altered to read, "Too many people are living in an unnatural, soft way; we should return to a more red-blooded active way of life."

Tendency to accept rational authority was measured by a test composed of the items shown in Table 1 (The RA scale). Both positive and negative responses were included in an effort to control response set. The items are based on the author's interpretation of Fromm's concept, as elaborated in the above-mentioned references. Only face validity is claimed.

The items of both scales were mixed with each other and with those of several other personality and attitude scales, making a booklet of 150 items. Subjects were instructed to respond *True, Cannot Say,* or *False.* Only *True* and *False* responses were scored.

TABLE 1

ACCEPTANCE OF RATIONAL AUTHORITY SCALE

1. When a person in authority tells you to do something, you should do it, even though you don't see a good reason for it, provided they show you it is for your own good. (True)
2. It is generally best to be a little suspicious and critical of persons in a position of authority, or they may get too much like dictators. (False)
3. Most people who are leaders in the world today got there by crooked or sneaky means. (False)
4. No man really has the right to give another orders unless he can back it up with force. (False)
5. Someone who knows more than you do should be treated with respect, even if you don't think much of him personally. (True)
6. There isn't really very much your parents or older people can tell you that will help you get along in the world nowadays. (False)
7. Most "experts" are really only trading on their reputations to let them lord over people. (False)
8. It's no wonder some teachers have trouble keeping order in schools; they try to be bossy and run things that should be left to the pupils. (False)
9. When someone has spent years studying a subject, you can usually be sure that he will be better at it than you will. (True)
10. All teachers should be treated with more respect than the average person. (True)
11. The world is full of snobs who put on airs but aren't really better than anyone else at doing most of the things in life. (False)
12. Most people who regard themselves as authorities or experts should be taken down a peg now and then. (False)
13. The best school system is one that is democratic and treats all the pupils exactly alike. (False)
14. Complete freedom is the best way to bring up a child if you want it to be free and active. (False)
15. Most so-called "juvenile delinquency" is really just "youthful exuberance" and should not be punished. (False)
16. It is a good idea for pupils to give teachers a hard time; it keeps the teacher from getting too conceited and "drunk with power." (False)
17. "Bohemians" and people who lead wild lives are just unwilling to accept the fact that some kinds of social controls are necessary (for example, laws concerning what you can wear on the street). (True)
18. One of the best attitudes a young person can learn is that "nothing is sacred." (False)
19. Most policemen are fair and honest. (True)

2. *Subjects*

Both scales were administered to 438 basic airmen who had entered the Air Force only a few weeks previously. Most were between the ages of 17 and 22, and were high school graduates.

B. RESULTS

1. *Measurement of Acceptance of Rational Authority*[2]

The practicality of regarding this characteristic as a measurable and unitary trait was tested by computing the RA test's split-half reliability. There were equal numbers of positively and negatively keyed items in each half. This statistic could be computed for only 190 of the subjects, but there is no reason to believe that they are unrepresentative of the rest. The split-half correlation was found to be .46 which, when corrected by the Spearman-Brown formula, yields a reliability estimate of .63. Both values are significant beyond the .01 level.

2. *Relationship to Acceptance of Irrational Authority*

The correlation (Pearson r) between the RA scale and the F-scale, with $N = 438$, was found to be $r = .04$. This value is not significant at the .05 level.[3]

C. DISCUSSION

From these results, we conclude, first, that it is possible to measure the tendency to accept rational authority as a unitary trait. Indeed, success with this short scale suggests that future refinement and lengthening could produce an instrument fully as reliable as the F-scale itself. Further research on this important variable appears possible.

Second, the lack of relationship between acceptance of the two types of authority suggests that considerable caution should be exercised both in theoretical speculations on the nature of the "authoritarian personality" and in practical recommendations for the sorts of attitudes one should attempt to develop in youngsters for optimal "mental health." With respect to the former, a negative relationship might have been expected, reasoning that one who democratically accepted rational authority would resist an autocratic, irrational authority. The absence of a relationship suggests caution in the

---

[2] No reliability coefficient was computed for the F-scale since extensive research by this author (Rudin, 1958) and others has obtained consistent reliability estimates of .85-.95.

[3] Other statistics of interest for normative purposes:

Test	Mean	Standard Deviation
RA	12.130	3.042
F	16.995	4.381

use of such broad categories as "authoritarian personality" and attention to the specific social context in which a belief is acquired and expressed.

With respect to the latter, it is common practice among some clinicians and some social, educational, and developmental psychologists to urge parents to exercise only slight authority over a child, adopting a permissive, "non-authoritarian" (*sic!*) attitude, in the belief that this will prevent the child's developing dependence on irrational authority. Their assumption that decrying respect for *all* authority, regardless of its nature, will decrease respect for irrational authority in particular, is not borne out by these data. Indeed, it seems probable that by *not* teaching respect for rational authority, the traditional teacher-student and parent-child relationship is broken, and immature youngsters, seeking as they are a source of norms, are forced to turn to the frequently irrational authority of the peer group.

D. SUMMARY

Research on attitudes toward authority has concentrated on the acceptance of irrational authority, as exemplified by the master-slave relationship and measured by instruments such as the F-scale. This research studied the tendency to accept rational authority as exemplified by the teacher-student relationship and measured by a scale devised by the author. The scale demonstrated reliability and face validity adequate for the purposes of this paper. It correlated essentially zero with the F-scale when given to a sample of 438 basic airmen. It was concluded that there is no necessary relationship between the tendency to accept irrational authority and the tendency to accept rational authority. By implication, it would seem that "permissive" methods of child-rearing may not only fail to protect the person against future submission to irrational authority, but might also disrupt normal relations with rational authority figures.

REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, E., LEVINSON, D. J., & SANFORD, R. N. The authoritarian personality. New York: Harper, 1950.
2. FROMM, E. Escape from freedom. New York: Rinehart, 1941.
3. ———. Man for himself. New York: Rinehart, 1947.
4. RUDIN, S. A. Academic anti-intellectualism as a problem in student counseling. *J. Counsel. Psychol.,* 1958, **5**, 18-23.

School of Psychology
Georgia Institute of Technology
Atlanta 13, Georgia

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 185-192.

DELAYED AUDITORY FEEDBACK AND ETHANOL: EFFECT ON VERBAL AND ARITHMETICAL PERFORMANCE*

Indiana University School of Medicine

ROBERT B. FORNEY AND FRANCIS W. HUGHES

A. INTRODUCTION

It has always been recognized that alcohol shifts patterns of behavior from those purported to be normal. This can appear in numerous guises. For example, with many persons alcohol unveils an aggressive if not pugnacious character. In others, there is elicited social withdrawal and a depressive diathesis. Such a diversity in reactivity causes great difficulty in arriving at a critical evaluation of behavioral degeneration while under the influence of alcohol. We have selected a particulate of behavior, viz., performance, in order to investigate the influence of alcohol on verbal and arithmetic scorings.

The tasks employed in this study proved to be extremely simple for our subjects. Rather than compounding the arduousness of single tasks, a distracting force was introduced. This consisted of a controlled delayed auditory feedback system. It has been our experience that such an impairment measurably reduces the speakers output, as well as accuracy or articulation of speech.

It is quite possible, if not likely, that ingestion of the same quantities of alcoholic beverages can cause quantitatively different impairments in performance in different subjects. It is our contention that data obtained for subjects receiving alcohol is reliable only if one determines and reports the actual concentration of ethanol within the blood stream of the subjects rather than the amount ingested. This was controlled in our experiment by testing all subjects at a time when their blood alcohol concentrations were relatively the same.

B. METHODS

1. *Subjects*

The subjects in this study were 7 male and 3 female paid volunteers. They knew nothing of the test procedure until 15 minutes before the test,

---

* Received in the Editorial Office on April 6, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

at which time they were given a brief description of the apparatus and procedure. They were urged to complete as much as possible of each test in order to gain a higher score but were cautioned that penalties would be imposed for articulatory as well as mathematical errors. The subjects were medical or graduate students.

2. *Procedure*

Each subject was tested in a soundproof room. Six different tests were administered. For each test, the subject spoke into a Sure 55S microphone which fed an Ampex 350 recorder operating at 7½ inches per second. The playback head (2 1/16 inches behind the recording head) was monitored with a 20 watt Heath amplifier into a set of headphones equipped with rubber cups. Thus an 0.28 second delay audio feedback was produced. The sound level was maintained higher than a comfortable listening level so that the subject could only hear the feedback.[1] The method of inducing disturbance in speech is a modification of those described by several other investigators. (1, 2)

The subjects received two series of tests. One was without any interfering drug and the other was after the ingestion of alcohol. Each subject consumed alcohol, as Vodka, mixed with a carbonated beverage. Twenty minutes after the completion of drinking, blood alcohol levels were determined from alveolar air analysis using the Borkenstein Breathalyzer (4). The subjects drank the beverages in quantities calculated to contain 45 ml. of 200 proof alcohol per 150 lb. body weight. In eight of the ten subjects, this effected a blood alcohol concentration of approximately 100 mgm. per cent. It was necessary that two subjects consume additional alcohol to reach this level. This concentration was selected because definite impairment in performance is reported in all individuals at a blood level of approximately 100 mgm. per cent (3).

3. *Tests*

There were six test forms:

Test one—Verbal Output. A passage from Aristotle's Selections (F. H. Peters, Fine Editions Press) was chosen. The selected section was irrelevant, in that it would have been essential for the reader to have read the preceding and continuing passages for complete comprehension. The participants were told to read as rapidly but distinctly as possible for a two minute period.

---

[1] We wish to thank Lt. Paul Lowe and the Indiana State Police for their assistance in the preparation and use of the audio feedback system for these tests.

Test two—Forward Count. Each subject was asked to count from one, forward, as rapidly as possible for two minutes. He was cautioned against slurring or short cuts. For example, he was to say "one hundred and twenty-three" rather than "one twenty-three."

Test three—Reverse Count. The starting number was five hundred. From this number the subject was to count backward as quickly as he could for a two minute period.

Test four—Progressive Count. Starting with one, the subject was to count in multiples of four, e.g., 1, 5, 9, 13, 17, etc. This test also lasted two minutes.

Test five—Addition. (Figure 1) This consisted of a mimeographed sheet which was to be read. These were simple arithmetical problems in addition. The subject read aloud all the numbers and supplied the calculated answer.

Test six—Subtraction. (Figure 2) This was similar to Test four except that simple mathematical problems in subtraction were used.

4. *Scoring*

Each test was timed carefully. Although the subject might pursue the task beyond the allotted time, only two minutes were actually recorded on tape. After the two minutes had passed, a buzzer notified the subject of the completion of this part of the test. Another buzzer, relayed from the adjacent monitoring room, indicated that he was to start the next test. The tapes were collected and scored later by an associate of the investigators. The grader of the tests did not know which of the tapes contained records of the subjects after drinking, but this soon became apparent. Figure 3 is a reproduction of the evaluation sheet used by the scorer.

C. RESULTS

The delayed auditory feedback system was sufficiently distracting and annoying to make a simple task difficult. Certain segments of performance, under ordinary circumstances, are automatic. With this distraction, our subjects had to concentrate, not only to perform the test, but to compensate for the frustration encountered by the difficulty in performing such simple tasks.

Mean alcohol levels, calculated as blood concentrations, were 97 mgm. per cent (S. E. 3mgm. per cent).

FIGURE 1

ADDITION TEST READ BY SUBJECTS DURING STIMULUS OF AUDITORY FEEDBACK SYSTEM

Name .... Date ....

Addition:

Complete the following: (Quickly as possible, start on signal, stop when buzzer sounds)

1. Two plus three equals ....
2. One plus four equals ....
3. Nine plus three equals ....
4. Eleven plus five equals ....
5. Seven plus eight equals ....
6. Nine plus one equals ....
7. Seven plus eleven equals ....
8. Eleven plus eighteen equals ....
9. One plus two equals ....
10. Four plus four equals ....
11. Nineteen plus six equals ....
12. Six plus three equals ....
13. Eight plus eleven equals ....
14. Sixteen plus six equals ....
15. Twenty-seven plus six equals ....
16. Seven plus six equals ....
17. Nine plus nineteen equals ....
18. Nine plus eighteen equals ....
19. Three plus twelve equals ....
20. Five plus twenty equals ....
21. Nineteen plus nineteen equals ....
22. Eleven plus twenty-three equals ....
23. Fourteen plus seventeen equals ....
24. One plus one plus one equals ....
25. Nine plus three plus six equals ....
26. Six plus two plus three equals ....
27. Twelve plus three plus six equals ....
28. Twenty-two plus twenty-three equals ....
29. Six plus two plus sixteen equals ....
30. Nineteen plus nine plus nine equals ....
31. Two plus five plus sixteen equals ....
32. Sixteen plus two plus twenty-one equals ....
33. Two plus two plus twenty-seven equals ....
34. Thirty-one plus twenty-two equals ....
35. Six plus seven plus eight equals ....
36. Two plus nine plus eight equals ....
37. Sixteen plus eighteen plus five equals ....
38. Five plus five plus forty-eight equals ....
39. Seventy-five plus nine plus eight equals ....
40. Three plus eight plus nineteen equals ....
41. Six plus eight plus twelve equals ....
42. Eleven plus two two plus six equals ....
43. Three plus five plus seventy-eight equals ....

The results of the scorings of the tests are in Table 1. There was a significant increase in errors in the reverse count test and a decrease in number of problems completed in the addition test. Nine of the ten subjects indicated that they "thought they did much better" on the series of tests after alcohol. The gross signs of frustration and nervousness were lessened. However, the mean scores show no improvement in any of the tests.

FIGURE 2

SUBTRACTION TEST READ BY SUBJECTS DURING STIMULUS OF AUDITORY FEEDBACK SYSTEM

Name .... Date ....

Subtraction:

As quickly as possible, complete the following. Say aloud all numbers involved.

1. Seven minus two equals ....
2. Three minus one equals ....
3. Nineteen minus twelve equals ....
4. Sixteen minus two equals ....
5. Thirteen minus twelve equals ....
6. Twelve minus two equals ....
7. Fourteen minus three equals ....
8. Ten minus seven equals ....
9. Fourteen minus seven equals ....
10. Sixteen minus eleven equals ....
11. Eleven minus nine equals ....
12. Six minus four equals ....
13. Fourteen minus three equals ....
14. Ten minus two equals ....
15. Twenty-seven minus fifteen equals ....
16. Fifteen minus eight equals ....
17. Twenty-one minus seven equals ....
18. Thirty-eight minus nine equals ....
19. Twenty-seven minus sixteen equals ....
20. Seven minus five equals ....
21. Thirteen minus two equals ....
22. Ten minus three equals ....
23. Forty-seven minus twenty-one equals ....
24. Twenty-one minus nineteen equals ....
25. Ninety-six minus five equals ....
26. Twelve minus eleven equals ....
27. Sixteen minus nine equals ....
28. Thirteen minus eight equals ....
29. Twenty-seven minus six minus one equals ....
30. Twenty-eight minus one minus six equals ....
31. Sixty minus forty equals ....
32. Forty minus nineteen minus nine equals ....
33. Nine minus three minus two equals ....
34. Sixty minus twenty-one minus twelve equals ....
35. Sixteen minus two minus nine equals ....
36. Forty minus eighteen minus three equals ....
37. Nineteen minus nine minus nine equals ....
38. Nine minus two minus three equals ....
39. Forty-seven minus sixteen minus twelve equals ....
40. Nineteen minus twelve minus two equals ....
41. Seventy-five minus two minus seventeen equals ....

FIGURE 3

EVALUATION SHEET

Name .... Age .... Wt. .... Sex .... Date ....

Medical History (Drugs, etc.) ....

Blood Alcohol .... before test

.... after test

Comments ....

....

Other Test Medication ....

- I. *Verbal Output*
  - No. Words per 2 minutes ....
  - Errors
    - omissions ....
    - insertations ....
    - mispronunciations ....
    - spasms (stammers) ....
    - prolongations ....
    - other ....
  - Total Errors ....
  - Total Verbal Output ....
  - % Errors ....
- II. *Forward Count*
  - Last Number ....
  - Total Numbers ....
  - Errors
    - omissions ....
    - mispronunciations ....
    - spasms (stammers) ....
    - prolongations ....
    - other ....
  - Total Errors ....
  - Forward Count Total ....
  - % Errors ....
- III. *Reverse Count*
  - Total Numbers Counted ....
  - Errors
    - omissions ....
    - mispronunciations ....
    - spasms (stammers) ....
    - prolongations ....
    - other ....
  - Total Errors ....
  - Reverse Count Total ....
  - % Errors ....
- IV. *Progressive Count*
  - Total Numbers ....
  - Errors
    - omissions ....
    - addition errors ....
    - mispronunciations ....
    - spasms (stammers) ....
    - prolongations ....
    - other ....
  - Total Errors ....
  - Progressive Count Total ....
  - % Errors ....
- V. *Addition*
  - Number Problems Completed ....
  - Errors
    - additions incorrect ....
    - problems misread ....
  - Total Errors ....
  - % Errors ....
- VI. *Subtraction*
  - Number Problems Completed ....
  - Errors
    - subtractions incorrect ....
    - problems misread ....
  - Total Errors ....
  - % Errors ....

D. DISCUSSION

Our data generally indicate that verbalization is not significantly impaired in subjects who had a blood alcohol concentration of approximately 100 mgm. per cent. Ethanol is a progressive depressant of the central nervous system. Subjective as well as objective symptoms verify this. Either increased verbalization or decreased efficiency of speech may occur after alcohol is ingested but these are not predictable or useful criteria of shades of inebriation.

TABLE 1
MEAN TOTAL SCORES AND PERCENTAGE ERRORS BEFORE AND AFTER INGESTION OF ALCOHOL*

	Verbal	Output	Forward	Count	Reverse	Count	Progressive	Count	Addition		Subtraction	
	Total Words	% Errors	Total Nos.	% Errors	Total Nos.	% Errors	Total Nos.	% Errors	Total No. Problems	% Errors	Total No. Problems	% Errors
Without Alcohol	313.7	7.2	151.2	10.3	87.5	9.3	68.8	24.9	30.7	14.5	33.4	11.9
After Alcohol	336.8	6.9	168.2	10.7	84.7	20.8	66.3	22.4	33.6	28.3	33.2	18.6
P Pre-drug vs. Alcohol	n.s.	n.s.	n.s.	n.s.	n.s.	<0.02	n.s.	n.s.	<0.05	<0.10 >0.05	n.s.	n.s.

---

* Mean Blood Alcohol Concentration = 97 mgm. per cent (S.E. ± 3 mgm. per cent)

Deficiency in verbalization occurs after introducing a distraction of auditory feedback but this was not enhanced by alcohol in our experiments.

In order to verbalize mathematical calculations, a component of performance which is less reflex in character than pure verbalization is necessary. Greater mental activity or concentration must be employed. Alcohol, in addition to the reverberation of the subject's own voice, was sufficiently distracting to increase the incorrectness of responses in several mathematical tests.

The specific purpose of this experiment was to investigate the usefulness of delayed auditory feedback system in the evaluation of centrally acting drugs. It is the intent of the investigators to modify the tests and further increase or decrease the demands on mental activity of the subjects. With such tests, we hope to explore the actions of such agents as tranquilizers, sedatives and psychostimulants on the responsiveness of subjects to delayed auditory feedback stimuli.

E. Summary and Conclusions

Ten adult subjects recorded verbal and arithmetical tests while distracted by an auditory feedback system. They were tested before and after ingesting a quantity of an alcoholic beverage calculated to effect a 100 mgm. per cent blood level of ethanol.

The following tentative conclusions are made:

1. Alcohol did not significantly increase scores in any of the tests.
2. Verbalization was unaffected.
3. Ethanol tended to interfere with test performances in which mathematics are involved. The arithmetical tests in which such deficiencies became evident involved simple addition exercises and reverse counting.

References

1. Black, J. W. The Effect of Delayed Side-Tone upon Vocal Rate and Intensity. *J. Speech Hearing Res.*, 1951, **16**, 56-60.
2. Chase, R. A., Sutton, S., & First, D. Bibliography; Delayed Auditory Feedback. *J. Speech Hearing Res.*, 1959, **2**, 193-200.
3. Harger, R. N. Proc. Symp. on Alcohol and Road Traffic. Dept. Police Admin., Indiana Univ., Bloomington, Indiana, 1958, p. 275.
4. Monnier, D. Un Nouvel Appareil de Depistage de L'Alcoolisme: Le Breathalyzer. *Extrait de la Rev. Internat. de Criminol. et de Police Technique,* 1956, **10,** 1-6.

Indiana University Medical Center
1100 West Michigan Street
Indianapolis 7, Indiana

Published as a separate and in *The Journal of Psychology*, 1961, 52, 193-200.

A MODIFIED VERSION OF THE WISCONSIN GENERAL TEST APPARATUS*[1]

Department of Psychology, Brown University

Allan M. Schrier

A. Introduction

This paper describes the version of the Wisconsin General Test Apparatus (WGTA) used to train monkeys at the Walter S. Hunter Laboratory of Psychology at Brown University. The WGTA was first described by Harlow and Bromer (1938). A brief description of a later, improved WGTA was published by Harlow (1949). In addition, Riopelle (1954) has published a general description of a modified version of the 1949 WGTA. The version of the WGTA described in detail below, though differing in several respects from the earlier versions (for example, size, number of screens, and mode of operation of the screens), retains their basic features.

B. The Apparatus

A photograph of the version of the WGTA used at Brown is shown in Figure 1, and three views of the apparatus drawn to scale are presented in Figures 2, 3, and 4. The drawings are rather schematic and are intended to provide an idea of the basic dimensions of the apparatus and the placement of some of the more important components. The apparatus consists of a restraining cage and table, the table situated immediately before the cage. The table supports a superstructure which holds three screens, a stimulus tray, and a light fixture. The *E* sits at the table facing the restraining cage. A writing stand is beside *E*.

The table and superstructure. An opaque screen, composed of black plastic, and an adjacent transparent plastic screen are located immediately in front of the restraining cage. Both of these screens are 22 in. high. A screen composed primarily of one-way vision glass is situated near *E*'s end of the table.

---

* Received in the Editorial Office on April 12, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The establishment of a primate colony by the author at the Walter S. Hunter Laboratory of Psychology and the construction of the apparatus described in this report were supported in part by a research grant (M-2818) from the National Institute of Mental Health, Public Health Service.

The opaque and transparent screens (the "forward" screens) are interposed between *S* and the stimulus tray between trials, whereas the one-way vision screen is interposed between *E* and the stimulus tray during a trial. The one-way vision screen is counterbalanced and controlled manually. The

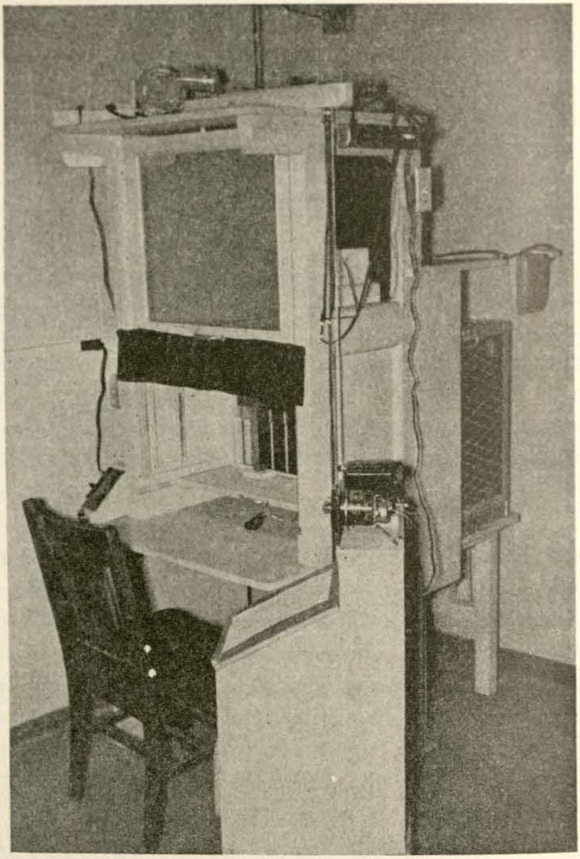


FIGURE 1
The test apparatus. Note some of the components not shown in the apparatus drawings—the canopy extending out over the restraining cage, the extension of the side of the superstructure, which serves in part to support the canopy, and the lever switches, controlling the forward screens, which can be seen, in the photograph, just above the top of the chair for *E*. The black cloth hanging from the bottom of the one-way vision screen was added because this screen cannot be lowered below the grooves for the stimulus tray.

forward screens are raised and lowered by means of motors, a Dayton reversible gear motor (Model 4K862 with a gear ratio of 238/1) for the opaque screen, and a more reliable Slo-Syn reversible synchronous motor (Model SS250-P1) for the transparent screen. The raising of the opaque screen and the lowering of the opaque and transparent screens are controlled by lever switches located near *E*. The transparent screen is automatically raised when the top of the opaque screen is raised past a point at which a Micro Switch is located, in most experiments, approximately 31 in. above

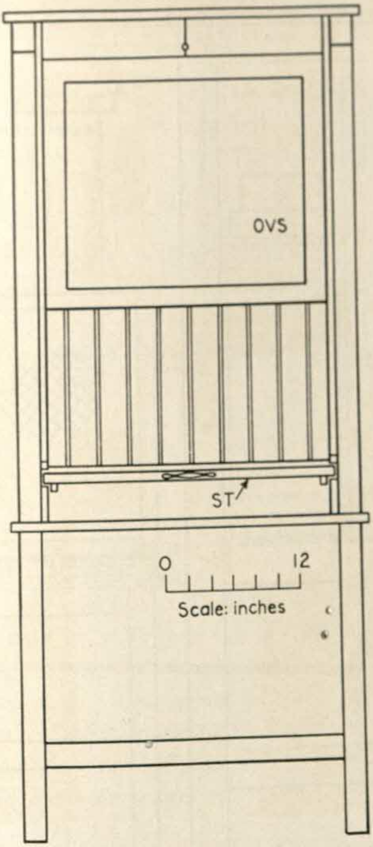


FIGURE 2
VIEW OF *E*'S END OF THE APPARATUS. The table supporting the restraining cage is not shown. ST: Stimulus tray. OVS: One-way vision screen.

the stimulus tray. The extent of the upward and downward movements of the forward screens is limited by means of appropriately placed Micro Switches. The forward screens are motor driven in order to increase the reliability of measurements of response latency, and also, to prevent the possibility, when discrimination problems are presented, of extra cues resulting from the raising of a screen at different rates depending on whether *S* is orienting toward the correct or incorrect stimulus.

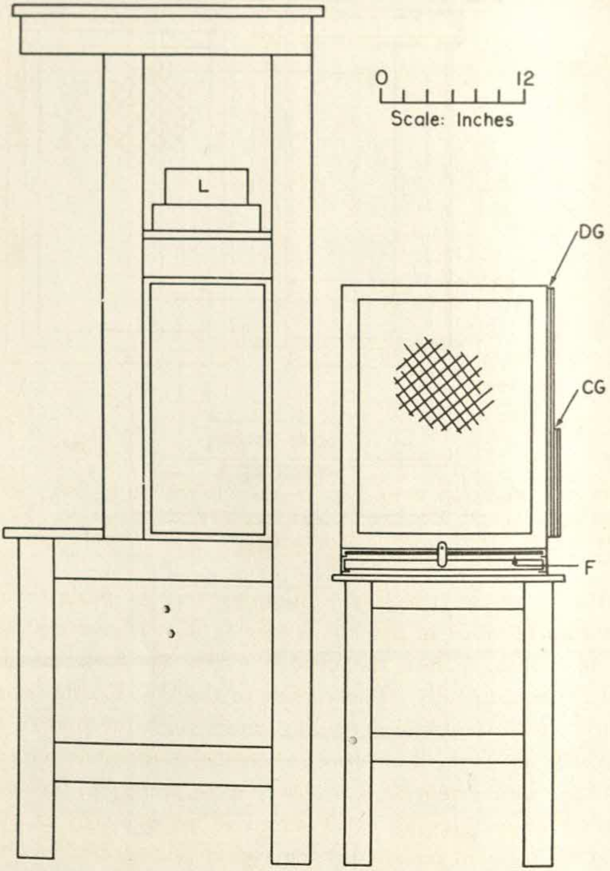


FIGURE 3
SIDE VIEW OF APPARATUS. L: Fluorescent light fixture. DG: Groove for restraining cage door. CG: Groove for transfer cage.

The use of a transparent screen followed from a decision to employ a test procedure in which the stimulus tray is pushed forward within reach of *S* before the opaque screen is raised, in contrast to the procedure, used in some laboratories, in which the stimulus tray is pushed forward as the opaque screen is raised or after it is raised. The transparent screen prevents *S* from responding before the opaque screen has been raised high enough to reveal the stimuli fully. The version of the WGTA described by Riopelle has a sheet of clear plastic extending down from the bottom of the opaque screen, which serves the same function as the transparent screen. In general, the use of a plastic extension is probably to be preferred, because it eliminates the necessity of working an extra screen either manually or by motor. The choice of two forward screens for the apparatus described here was motivated in part by a desire to keep the visual field above the stimulus tray perfectly clear for experiments involving vertical separation of the relevant cues from the site of the instrumental response and the reward. It appeared that the plastic extension used by Riopelle would remain interposed between *S* and stimulus objects presented several inches or more above the level of the stimulus tray, unless the height of the superstructure were increased, thus

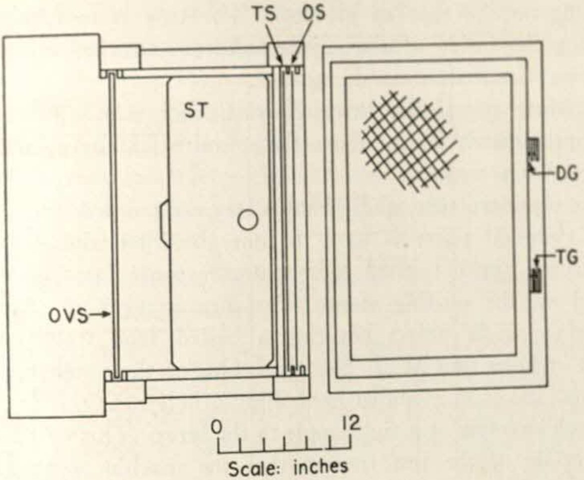


FIGURE 4
TOP VIEW OF APPARATUS WITH THE SUPERSTRUCTURE CUT AWAY AT THE LEVEL OF THE FLUORESCENT LIGHT FIXTURE. OVS: One-way vision screen. ST: Stimulus tray. TS: Transparent screen. OS: Opaque screen. DG: Groove for restraining cage door. TG: Groove for transfer cage.

increasing the height to which the plastic extension could be raised. Practical limitations on the height of the apparatus, among other things, mitigated against the use of the extension.

The stimulus tray contains one or more food wells, the number and their size depending on the type of problem. The tray is mounted in grooves running along the inside of the superstructure, and can readily be moved toward or away from the restraining cage.

A light fixture containing three 20-watt, cool white fluorescent lamps is located approximately 21 in. above the stimulus tray between the forward and one-way vision screens.

The table, superstructure, and stimulus tray are painted gray.

An ATC interval timer is used to time the inter-trial interval, and a Standard electric timer is used to measure response latency. Both timers are mounted on the writing stand. For measurement of latency of displacement of a single object covering a baited food well, the Standard timer circuit includes two Micro Switches. One of these switches has a lever actuator which has been modified by drilling a hole near the tip of the lever to hold a machine screw at a right angle to the lever. This switch is mounted on the underside of the test tray so that the machine screw projects up through a small hole immediately in front of the food well. Placement of an object over the food well closes this switch. The second Micro Switch has a long plunger actuator and is mounted on the underside of the table so that the plunger projects up through a hole in the table directly beneath the transparent screen. The release of the plunger when the transparent screen is raised completes the timer circuit through the normally-closed circuit of the second Micro Switch. Displacement of the object by *S* breaks the timer circuit. A hold circuit added to the timer circuit prevents *S* from reactivating the timer when reaching into the food well. The depression of the plunger actuator when the transparent screen is lowered results in the activation of the ATC interval timer through the normally-open circuit of the Micro Switch. The closing of a switch in the ATC timer at the end of the inter-trial interval automatically resets the Standard timer to zero. The noise of the resetting mechanism provides *E* with a convenient cue for the start of a new trial.

The restraining cage. The restraining cage is constructed of galvanized iron and rests on a wooden table. The top, floor, and three sides were originally composed of ¾-inch diamond mesh. The side facing the table consists of a series of ⅜-inch vertical rods centered about 2½-inches apart.

The floor of the restraining cage is removable. Just below the floor is a stainless steel waste pan.

A number of steps were taken to make the restraining cage environment more homogeneous and to reduce the possibility of artifacts introduced by changing shadow patterns on the walls and ceiling of the test room as *E* prepared the stimulus tray between trials. The mesh sides of the cage were covered with aluminum sheeting painted flat black. A canopy extending out from the superstructure over the top of the restraining cage was added. The sides of the superstructure were extended in order to block *S*'s view of the test room between the restraining cage and the superstructure as well as to provide support for the canopy.

A light fixture housing a 60-watt, warm white incandescent lamp is located in the center of the canopy.

White noise is supplied through three speakers, one mounted on the superstructure above the canopy, and one mounted on each side of the restraining cage between the canopy and the top of the restraining cage. The noise level at the center of the cage is approximately 82 db (0 db $= 0.0002$ dyne/cm^2).

It will be noted that two adjacent grooves are located on either side of the doorway of the restraining cage (see Figures 3 and 4). The inner grooves hold the sliding door of the restraining cage in place. The outer grooves are used to hold the transfer cage opposite the door of the restraining cage. Each of the transfer cages has a pair of flanges at one end, which extend out at right angles, and which fit into the grooves of the restraining cage. The home cages were built to receive the transfer cages in the same manner.

C. The Test Procedure

The usual test procedure is as follows: With the forward screens down, the stimulus tray is prepared and pushed forward. The one-way vision screen is lowered, and then the opaque screen is raised, followed by the automatic raising of the transparent screen. After *S* responds, the transparent screen is lowered, and then the opaque screen. The one-way vision screen is raised, the result of the trial recorded, and the stimulus tray pulled back and prepared for the next trial.

References

1. Harlow, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51-65.
2. Harlow, H. F., & Bromer, J. A. A test-apparatus for monkeys. *Psychol. Rec.*, 1938, **2**, 434-436.
3. Riopelle, A. J. Facilities of the Emory University Primate Behavior Laboratory. *J. Psychol.*, 1954, **38**, 331-338.

Department of Psychology
Brown University
Providence 12, Rhode Island

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 201-210.

MEANINGFULNESS (*m*) AND TRANSFER PHENOMENA IN SERIAL VERBAL LEARNING*

Department of Psychology, Montana State University

CLYDE E. NOBLE[1]

A. PROBLEM

This experiment is part of a series designed to identify and quantify the variables influencing performance in verbal learning. One powerful factor in the rote learning of words and paralogs is *meaningfulness* (*m*), defined by the operations of the association-frequency method (9, 13). Several studies have shown that manipulation of *m* facilitates the acquisition of skill in both serial learning (6, 10) and paired-associate learning (4, 12), but the transfer properties of materials varying in *m* value are unknown. The purpose of the present investigation was to determine whether there are any unequal transfer-of-training effects among lists differing widely in difficulty due to meaningfulness. Given equal amounts of practice and training, is inter-list transfer dependent on the levels of meaningfulness used in training?

Phrased in this way, the question becomes part of the more general problem of discovering the task factors which affect transfer between learning situations. According to one of Thorndike's (14) neglected—but not yet outmoded—principles, amount of transfer will depend on the proportion of common elements shared by the tasks. In the field of motor skill Bilodeau (3) found that, although the target size of a two-hand coordination apparatus influenced difficulty in *E*-paced following pursuit, there was no differential transfer from large to small targets as compared with changing from small to large targets. On the other hand, when the number and variety of responses were manipulated in a pair of *S*-paced recessed-tracking devices,

---

* Received in the Editorial Office on April 18, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This work was supported in part by a contract between the Office of Naval Research and Montana State University. Reproduction in whole or in part is permitted for any purpose of the United States Government. Assisting in data collection and statistical analyses were Blaine L. Baker, Ridgely W. Chambers, James E. Fuchs, George V. C. Parker, Donald P. Robel, and Richard K. Smith. An abridged report was made to the Midwestern Psychological Association, May 1959.

Jones and Bilodeau (7) did obtain differential transfer. The positive transfer effect produced by changing from a complex task (clover pattern) to a simple task (circle pattern) was not only greater than that resulting from the reverse order of events, but was as great as the effect of practicing directly on the simple task. Presumably, the efficacy of scheduling the complex task first resulted from the fact that it included the S-R components of the simple task (e.g., the clover track required three times as many directional shifts as the circle); this provided varied practice on a larger proportion of criterion responses. A rotary pursuit study of target-rate variations by Ammons, Ammons, and Morgan (1) showed that transfer effects were non-monotonic functions of the degree of difference between the first and second tasks.[2] This finding underscores the importance of similarity relationships, but since amount of practice co-varied with target speed these data are not as relevant as the two-hand coordination work.

What are the implications of the motor skill experiments for verbal learning? To answer this question, one must ask another: Does a difficult list of low *m* value (L) contain the S-R elements of an easy list of high *m* value (H)? Assuming Thorndike's hypothesis is valid, the reader's answer should probably be in the negative because of the total lack of communality between lists constructed from the extremes of the *m* scale. An affirmative answer would presuppose some kind of unequal, non-specific learning-to-learn transfer process which operates to facilitate the L→H sequence more than the H→L sequence. Of course, it is conceivable that different "general principles, modes of attack, and sets to perform" (8, p. 327) might be acquired under L-training and H-training, and that prior experience with such diverse materials could produce significantly different transfer effects. Andreas reminds us that "performance on a second task may be benefited by an earlier activity even when there are no mutual identical elements" (2, pp. 471-472). A third view is often held by educational, industrial, and military consultants who recommend an easy→difficult sequence for nearly every transfer situation they encounter. Perhaps such "experts" are unaware that "simple-easy" and "complex-difficult" are not necessarily synonymous; that the antecedent terms of each pair are *task* descriptions while the consequent terms are *behavioral* descriptions (*cf.* 11). The point is that difficulty *per se* (i.e., low proficiency scores) is an effect, not a cause, even in the more obvious part→whole transfer situations.

Summing up the evidence on the transfer-difficulty issue, Deese says that "in cases in which difficult problems involve stimulus-response components identical to those in easier tasks (but also additional stimulus-response components) there will be more transfer from the difficult problem than from the easier one. In those cases in which the problems involve the isolation of relevant stimulus dimensions and the association of certain stimuli with reward or solution of a problem, training on easy problems will transfer more than training on difficult problems" (5, p. 231). Since the present experiment seems to fit neither of Deese's examples, the Thorndikean theory disinclined us toward predicting any differential transfer effects. This presents a familiar statistical dilemma. As McGeoch and Irion put it, "to demonstrate the occurrence of zero transfer would require proof of the null hypothesis, a feat which experiment cannot accomplish" (8, p. 299). They add, however, that if theoretical analysis predicts no transfer, then "practically one may infer that none exists when every reasonable effort has been made to demonstrate transfer without result" (p. 300). Although the present experiment may be regarded as a "reasonable effort" in this direction, the major purpose of its design was to determine what transfer phenomena attend variation of the meaningfulness factor, and to measure their effects in serial learning. Knowledge of the transfer properties of different verbal materials is important for the proper design of future experiments in this program.

---

[2] A summary by the same authors appears in *Percept. Mot. Skills,* 1956, 6, 43.

B. METHOD

1. *Apparatus*

The learning device was a Stoelting memory drum which presented the items serially, the inter-stimulus and inter-trial periods being 2 and 4 sec., respectively. There were five serial lists of eight pronounceable two-syllable items each. A Practice list was selected from the central portion of the *m* scale (9) such that the average meaningfulness ($\overline{m}_M$) was 3.52. This list, being of median *m* value, was designated by the symbol M. Two lists of low meaningfulness ($\overline{m}_L = 0.31$) labeled L_1 and L_2, and two lists of high meaningfulness ($\overline{m}_H = 7.22$) labeled H_1 and H_2 were prepared for the Training and Transfer phases, respectively. List M and the matched pairs of Lists L and H are reproduced in Table 1, with $\overline{m}$ values listed for both the 1950 military (9) and the 1960 college (13) norms.

2. *Subjects and Instructions*

The *S*s were 120 undergraduates obtained from psychology courses at

TABLE 1

Serial Lists for the Three Phases of the Experiment With Mean Meaningfulness ($\overline{m}$) Calculated From Military (9) and College (13) Norms

Practice	Training		Transfer	
M	L_1	H_1	L_2	H_2
Mallet	Tarop	Army	Delpin	Region
Quota	Xylem	Jewel	Latuk	Insect
Yeoman	Bodkin	Office	Gokem	Wagon
Zenith	Nares	Village	Tumbril	Money
Pigment	Volvap	Garment	Icon	Jelly
Jitney	Ulna	Dinner	Sagrole	Captain
Rampart	Rennet	Heaven	Polef	Zebra
Argon	Kupod	Uncle	Balap	Kitchen
$\overline{m}_9 = 3.52$	0.31	7.19	0.31	7.25
$\overline{m}_{13} = 6.32$	3.70	9.64	3.36	9.59

Montana State University in 1958. There were 36 women and 84 men, whose ages ranged from 17 to 34 (mean = 20.6 years). All were naive to verbal learning. They were instructed as follows:

> "This is a test of your ability to learn a list of two-syllable words. Shortly after the apparatus starts, you will see an asterisk, then a word will appear in the window. You are to pronounce each word aloud as you see it. After you have seen the list once, your job will be to anticipate each word in the series. In other words, as you see one item you are to pronounce the one that comes next before it appears. The asterisk will warn you that the list is about to begin again.
>
> "Whenever you fail to anticipate a word, say it aloud when it appears anyway. If you think you know what the next item is, but are not sure, make a guess. It will not hurt your score, and if you get it right it counts as a success. If your guess is wrong, correct yourself aloud as soon as the word appears. Always pronounce the word distinctly. Do not try to use any special system in your learning. Simply associate each item with the next one as the series moves along. Any questions?
>
> "First we will have a practice list to acquaint you with the procedure; however, I will record your scores." (*E* gave 10 trials of List M.) "Now we'll have a short rest period." (*E* talked to *S* for 2 minutes, then gave 10 trials on the Training List: H_1 or L_1.) "We'll now have another short rest period." (*E* talked to *S* for 2 minutes, then either:) "Now let's continue with the same list" (followed by 20 more trials for Groups H_1-H_1 and L_1-L_1) or "Now let's learn a different list" (followed by 20 more trials for the four switched groups). "Please do not discuss the experiment with anyone. Thank you."

3. *Procedure*

The experiment was divided into three phases: Practice, Training, and Transfer. All *S*s began the Practice phase by receiving 10 trials on List M using the anticipation-correction procedure. After a 2-minute rest, the sample was divided into two matched groups of 60 *S*s each for the Training phase. They were trained on either List H_1 or List L_1 for 10 more trials. Following another 2-minute rest, the two Training groups were further divided into three matched subgroups of 20 *S*s each for the Transfer phase. Two subgroups from each Training condition were then transferred to a different list of either the same or opposite *m* value, while the remaining subgroup resumed practice as a control. All *S*s received 20 trials during the Transfer phase.

On the first presentation of each list (Trial 1) *S* merely read the item (word or paralog) aloud and standard pronunciations were established. Attempted anticipation of the correct responses ($R+$) began on the second presentation (Trial 2) of each phase, except for continuation Groups L_1-L_1 and H_1-H_1 during Transfer which *E* scored from Trial 1.

C. RESULTS AND DISCUSSION

General results of the experiment are shown in Figure 1. The data are expressed in terms of the mean percentage of correct responses (*R per cent*) as functions of the number of trials (N) for the various experimental groups during the Practice, Training, and Transfer periods. It will be convenient to discuss the results for each phase separately.

1. *Practice Phase*

Frequency scores of correct responses ($R+$) recorded from the 120 *S*s during the Practice period were arranged in a 9×6 mixed-factorial analysis-of-variance design to test for trend and level differences among the six subgroups ($n = 20$) in learning List M. The main effect of Trials was significant ($F = 150.48$; $df = 8/912$; $P < .001$), but there were no differences among Groups ($F < 1.00$; $df = 5/114$) and no Trials × Groups interaction ($F < 1.00$; $df = 40/912$). The latter two non-significant ratios indicate that the six subgroups to be formed in the Transfer period were quite comparable in ability and achievement before entering the Training phase.

2. *Training Phase*

Acquisition scores ($R+$) of the two matched groups learning Lists H_1

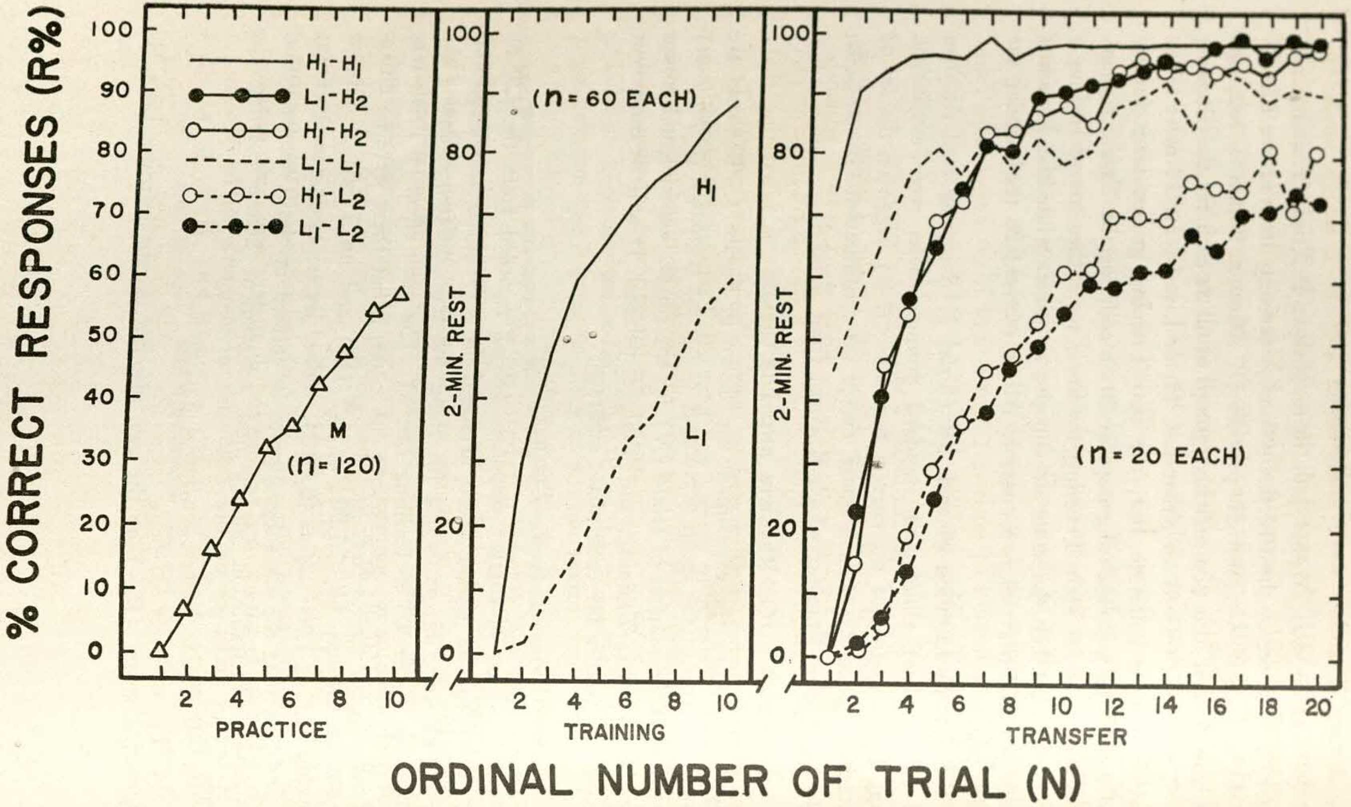


FIGURE 1
Acquisition curves for the three phases of the experiment. The symbols L, M, and H denote Low, Medium, and High meaningfulness (*m*) values. Each serial list contained 8 dissyllables, and the functions are based on 120, 60, and 20 *Ss* in the Practice, Training, and Transfer phases, respectively.

and L_1 during Training were analyzed in a 9 × 2 design similar to the one above ($n = 60$). As expected, the main effect of Trials ($F = 187.62$; $df = 8/944$; $P < .001$), the main effect of Meaningfulness ($F = 162.42$; $df = 1/118$; $P < .001$), and the Trials × Meaningfulness interaction ($F = 4.00$; $df = 8/944$; $P < .001$) were all significant. This analysis confirms the meaning-learning relationships reported earlier by Noble (10) and by Dowling and Braun (6). By Trial 10 of the Training period, Group H_1 earned an average *R per cent* score of 88.5 per cent while Group L_1 made only 59.8 per cent correct responses. Facilitation due to learning-to-learn is also evident in comparing the acquisition rate on List L_1 with the earlier performance on List M. Since the curves are very similar in slope, a substantial amount of positive transfer must have carried over from the Practice to the Training periods.

3. *Transfer Phase*

On the first post-rest trial of the transition to the Transfer period, Subgroups H_1-H_1 and L_1-L_1 both "forget" about 1.35 correct responses. These 17 per cent average losses are attributable to warm-up requirements and proactive interference from the two preceding lists. It is interesting that the amount of decrement is unrelated to the final proficiency level on Trial 10 of Training. Relearning during the next 20 Transfer trials continues at significantly different rates for the two subgroups, as found in Training, and the inter-trial variability is greater for the list of lower m value. Main effects of Trials ($F = 22.44$; $df = 19/722$), Meaningfulness ($F = 30.71$; $df = 1/38$), and their interaction ($F = 4.97$; $df = 19/722$) were all significant ($P < .001$).

Considering next the *S*s who transferred to List H_2, Subgroups L_1-H_2 and H_1-H_2 drop to zero on Trial 1 due to having new material. Thereafter they acquire proficiency at about the same rate as was true of List H_1 during the Training phase. This suggestion of non-differential transfer to lists of high m value is supported by a 19 × 2 analysis of variance which revealed significant main effects of Trials ($F = 100.60$; $df = 18/684$; $P < .001$), but no differences due to Groups ($F < 1.00$; $df = 1/38$) or the Trials × Groups interaction ($F < 1.00$; $df = 18/684$). The similarity of the H_1 and the H_2 curves in Training and Transfer means that learning-to-learn is either complete by the third list, or counteracted by proactive interference. Both H_2 curves are merging with the H_1-H_1 curve by the end of the Transfer period, being generally above the L_1-L_1 curve after Trial 8.

Finally, there are the *S*s who transferred to the L_2 list. Subgroups H_1-L_2 and L_1-L_2 fall to zero on Trial 1 for the same reason that the H_2 *S*s did, and they also improve at about the same rate that held for List L_1 in Training. It appears that Subgroup H_1-L_2 gains at a somewhat faster rate, but only the main effects of Trials are significant ($F = 99.85$; $df = 18/684$; $P < .001$). Neither the difference due to Groups ($F < 1.00$; $df = 1/38$) nor the Trials × Groups interaction ($F < 1.00$; $df = 18/684$) approaches significance. Although this 19 × 2 two-factor analysis failed to reveal any differential transfer to lists of low *m* value, a 19 × 2 × 2 three-factor Trials × Training × Transfer analysis was performed in an effort to detect possible higher-order interactions. Recalling the quotation from McGeoch and Irion (8) cited earlier, this test may be regarded as one which gives the null hypothesis another opportunity to be overthrown. If there really is greater positive transfer from high to low *m* value than from low to low, and if the gain is a function of Trials (as Figure 1 suggests), then such phenomena should at least be revealed as significant two-factor (Training × Transfer) and three-factor (Trials × Training × Transfer) interactions.

The summary in Table 2 indicates significant main effects of Trials and of Transfer Meaningfulness, but the only significant interaction is the familiar non-parallelism caused by differential *m* value when the pooled H_2 and L_2 curves are plotted against Trials. This interaction echoes the

TABLE 2
Analysis of Variance of Correct Responses ($R+$) Made by Subgroups L_1-L_2, H_1-L_2, L_1-H_2, and H_1-H_2 During Trials 2-20 of the Transfer Phase

Source	*df*	*MS*	*F*
Between *S*s	79		
Transfer Meaningfulness (B)	1	2273.61	65.86*
Training Meaningfulness (C)	1	9.01	<1.00
B × C Interaction	1	33.79	<1.00
Error (*b*)	76	34.52	
Within *S*s	1440		
Trials (A)	18	264.16	194.24*
A × B Interaction	18	6.84	5.03*
A × C Interaction	18	1.08	<1.00
A × B × C Interaction	18	.65	<1.00
Error (*w*)	1368	1.36	
Total	1519		

*$P < .001$

behavioral comparisons of the H_1 vs. L_1 curves in Training, and the H_1-H_1 vs. L_1-L_1 curves in Transfer. A multiplicative learning × meaningfulness relationship appears so consistently throughout experiments in this field that it is definitely attaining the status of an empirical law. It is concluded, therefore, that the operation of increasing *m* value is a temporary *performance* factor which powerfully facilitates rate of acquisition. That meaningfulness produces no permanent *learning* changes is implied by the absence of significant differential transfer-of-training effects.

D. Summary

A verbal learning experiment was conducted to discover transfer-of-training effects among 8-item serial lists differing in meaningfulness (*m*). For Practice, 120 *S*s received 10 trials on a list of average *m* value. The sample was then divided into two matched groups of 60 *S*s for Training on a list of either high (H_1) or low (L_1) *m* value for another 10 trials. Finally, six matched subgroups of 20 *S*s were formed for the Transfer phase of 20 trials. Two subgroups from each Training condition were changed to a different list of either the same or opposite *m* value, while the remaining subgroups resumed learning as controls.

The Training phase produced significant differences due to Trials, Meaningfulness, and their interaction, with List H_1 superior to List L_1. In Transfer, Subgroups H_1-H_1 and L_1-L_1 suffered equal decrements on the first post-rest trial, then continued relearning at significantly different rates consistent with the Training data. Subgroups L_1-H_2 and H_1-H_2 performed at the same rate as Group H_1 during Training, with no inter-group differences. Subgroups H_1-L_2 and L_1-L_2 resembled Group L_1 earlier, and likewise failed to show significant Transfer differences.

Given equal amounts of practice and training, inter-list transfer effects are independent of original *m* value. Our hypothesis that lists of high and low *m* value share no common elements was sustained by the absence of differential transfer. The data thus suggest a new principle of verbal learning and performance: meaningfulness facilitates rate of acquisition but has no influence upon transfer of training.

References

1. Ammons, R. B., Ammons, C. H., & Morgan, R. L. Transfer of training in a simple motor skill along the speed dimension. *USAF, Wright Air Development Center, Tech. Rep.*, 1954, No. 53-498.
2. Andreas, B. G. Experimental Psychology. New York: Wiley, 1960.

3. Bilodeau, E. A. A preliminary study of the effects of reporting goals as a function of different degrees of response accuracy. *USAF, Human Resources Research Center, Res. Bull.,* 1952, No. 52-4.
4. Cieutat, V. J., Stockwell, F. E., & Noble, C. E. The interaction of ability and amount of practice with stimulus and response meaningfulness (*m, m'*) in paired-associate learning. *J. Exp. Psychol.,* 1958, **56**, 193-202.
5. Deese, J. The Psychology of Learning. (2nd Ed.), New York: McGraw-Hill, 1958.
6. Dowling, R. M., & Braun, H. W. Retention and meaningfulness of material. *J. Exp. Psychol.,* 1957, **54**, 213-217.
7. Jones, E. I., & Bilodeau, E. A. Differential transfer of training between motor tasks of different difficulty. *USAF, Human Resources Research Center, Res. Bull.,* 1952, No. 52-35.
8. McGeoch, J. A., & Irion, A. L. The Psychology of Human Learning. (2nd Ed.), New York: Longmans, 1952.
9. Noble, C. E. An analysis of meaning. *Psychol. Rev.,* 1952, **59**, 421-430.
10. ———. The role of stimulus meaning (*m*) in serial verbal learning. *J. Exp. Psychol.,* 1952, **43**, 437-446; 1952, **44**, 465.
11. ———. Human trial-and-error learning. *Psychol. Rep.,* 1957, **3**, 377-398.
12. Noble, C. E., & McNeely, D. A. The role of meaningfulness (*m*) in paired-associate verbal learning. *J. Exp. Psychol.,* 1957, **53**, 16-22.
13. Noble, C. E., & Parker, G. V. C. The Montana scale of meaningfulness (*m*). *Psychol. Rep.,* 1960, **7**, 325-331.
14. Thorndike, E. L. The Psychology of Learning. New York: Teachers College, Columbia Univ., 1914.

Department of Psychology
Montana State University
Missoula, Montana

Published as a separate and in *The Journal of Psychology,* 1961, **52**, 211-222.

TOWARD CUTANEOUS ELECTRO-PULSE COMMUNICATION*[1]

Department of Psychology, Carnegie Institute of Technology

B. VON HALLER GILMER

A. INTRODUCTION

When consideration is given to the possibilities of utilizing the skin as a communication channel, several sets of questions have to be faced: What are the demands on the system in terms of information to be conveyed? Is the skin capable of transmitting information accurately, and under what conditions? Can cutaneous signals be coded to give both content and context to a language? Is such communication practical for complex information processing? These questions in turn lead to inquiries about the anatomy of the skin, neural coding, and about psychophysical functions.

There are at present two phases to our program of research in electropulse communication. One is the psychophysical approach to determine the boundaries between pain and painless pulses—the conditions and effects of intensity, polarity, duration, interval, electrode type, size, and spacing. All the conditions for pain-free stimulation commensurate with the demands of an easily learned coding system for a cutaneous language practical for the use of not only the normal, but also the deaf-blind person. Our second problem involves coding and complex information processing.

In developing the research we had considered concentrating our efforts entirely on the psychophysical problems in order to learn all the interrelationships between the many stimulus dimensions; but the problems here are endless, and we decided to restrict our efforts, to attack selected psychophysical and coding problems at the same time. We are asking the questions: What do we need to know to use the skin as a supplementary channel for information to the eye and ear? When used as a primary channel, what do we need to know for appropriate coding of a language for the deaf-blind, bearing in mind that they must communicate with people who may

---

* Received in the Editorial Office on April 26, 1961 and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research is being supported in part by a research grant, B-2022, from the National Institute of Neurological Diseases and Blindness.

not be motivated to learn some difficult or radically different new code? In other words, we are asking what psychophysical information we need, rather than attempting to learn all about skin sensitivity for its own sake.

B. Attempts at Cutaneous Communication

Among the first practical efforts to take advantage of the literacy of the human integument was that of Braille in 1826. The blind have been reading through their finger-tips with varying degrees of success for over a century, but the mastery of this language requires long practice as pointed out by Smith (22). Even when thoroughly mastered, Braille is still slow. Its efficiency is also reduced by such factors as excess perspiration and cold fingers. The maximum speed of Braille is considerably less than the cognitive processing time required, and this has caused some difficulties in use of the system.

Gault and his associates (4, 5) were another group who attempted, without success, to utilize the skin for purposes of communication. In essence these investigators tried to by-pass the ear by transducing speech sound into mechanical vibrations and apply them directly to the skin. Weiner, Wiesner, David, and Levine (27) more recently tried using alternating currents applied to the skin through a filter network of seven channels, but found the system not feasible for communication.

Although the possibilities of making use of the skin as another sensory communication channel have captured the imagination of investigators from time to time since the early basic works of von Frey and Kiesow (25), it was not until 1957 that the first successful attempt in tactile literacy was reported by Geldard (6). In 1960 fourteen investigators conferred at Fort Knox and opened up the basic problem areas of "asking the skin what language it could compass." Papers presented at this symposium have been edited by Hawkes (14) and represent the latest composite picture of the problems related to the development of a system for cutaneous communication. Geldard (7) states the over-all situation briefly when he says that "applied" problems in this area are themselves "basic" because so little is really known about the skin as a sense channel.

We know from the work of the Virginia group that vibro-tactile communication under laboratory conditions is possible (6). Subjects trained for less than 70 hours in this "vibratese" language were able to receive short words and messages at the rate of 38 words per minute with better than 85 per cent accuracy. Mechanical vibrations, however, have considerable drawbacks for communication through the skin because of the bulky nature of the transducers and of the diffuse spreading out of the mechanical stimulations. Electro-pulses, induced from square-waves, feel like slight "taps." They have an advantage over mechanical vibrations because they can be confined to a well-defined skin area. A disadvantage of the electro-pulses lies in the fact that they can induce pain under a number of conditions.

Although there have been many studies dealing with the measurement of the sensitivity of the skin to both mechanical vibrations and to alternating currents, the use of square-wave stimulations are of recent origin. In 1953 Sigel (21) demonstrated that within certain parameters, square-wave pulses delivered to the skin of the upper volar forearm resulted in pain-free perceptions of touch. He was interested in the clinical aspects of the stimulations. Hahn (13) in 1958 elaborated on this technique by using a square-wave source of high impedance. His primary finding was that threshold values vary primarily as a function of the duration of the square-wave. Gibson (8) has made use of pairs of biphasic electrical pulses in his attempts at producing painless taps.

C. Newer Possibilities of Cutaneous Communication

It looks as though it will be possible through the use of well defined electric pulses to develop a communication system providing information with such diversifications as stock market quotations, weather forecasts, and gun-laying data. From "vibratory barographs" for frogmen, to "covert codes" for use in secret transmissions, vibro- or electro-tactile communication falls into a minimum of seven classes of information. Through "amounts" one can present quantitative information, and through "coordinates" one can give relational information. In such a task as landing an aircraft, "directions" and "rates" can be transmitted through the skin. It has been shown that through vibratese one has the potential of a practical "language" system superior to the Morse Code. The attention-demanding qualities of a cutaneous signal give it some uniqueness in "warning," and possibly it may fit into some system involving "vigilance."

Between the eye and the ear, the ear is best in the temporal dimension, the eye in the spatial sense. The skin may be unique in combining the spatial-temporal dimensions in one sensory channel.

Much work has been done in identifying the stimulus dimensions for mechanical vibratory stimulation, but relatively little is known of the dimensions involved in electrical stimulation. Coding has had little work done on it. For one thing, it is difficult to decide whether attempts should be made to build up a general skin language, or coding should be made specific to each problem situation. What is the capacity of the observer in receiving information per unit of time? What is the capacity of the skin as a channel under different conditions? Gregg (12) has noted in studying some aspects of the air defense network that engineering the language instead of the machines has proved to be an effective way of overcoming the difficulties in the communication system. One may, then, postulate that by its very nature the skin is not handicapped with a lot of excess verbiage as is the written and spoken word. Perhaps the skin has possibilities of codes even superior to other channels because of its "simplicity."

Knowledge of stimulus variables in tactile sensitivity is almost non-existent. We need to learn more about skin input capacities, channel loads, the effects of distraction, and error ranges. The skin as a sensory channel may have one completely unique aspect; it is rarely ever "busy." True, it receives bumps, gets goose pimples, and wrinkles under water, but it is not used like the eyes and ears. This gives it an opportunity to learn, to become habituated to code that cannot be shut out; the skin cannot shut its eyes, it can hardly even holds its ears—in this respect the skin is a channel much more like the ear than the eye. Basic data needed may well be in signal-to-noise ratios for the skin.

The placing of electrodes is certain to be of importance in stimulating the skin in more than one place. And it appears that in setting up a language system one of the really important aspects is one of the "several eyes" or "several ears" kind of consideration. A big advantage of the skin is that it has many of these "severals" if we can find conditions under which the different skin areas can be made separate channels for information input.

Why should attention be given to building a communication system based on electrical impulse signals, particularly since so little is known about electro-anatomy? We have little or no knowledge about what kind of changes electrical inputs into the skin induce under conditions of mild stimulation. Gilmer and Haythorne (9) have shown that excessive stimulation of the skin with sinusoidal current can, under certain conditions, destroy epidermal and dermal tissue. What are the limits between "mild" and "excessive" stimulation that may be of importance in practical situations?

The recent works of Hahn (13), von Békésy (24), Gibson (8), Uttal (26), and Hawkes and Warm (15), have given encouragement to determine the conditions under which the skin can be stimulated electrically without the arousal of pain. Important to this problem are certain anatomical considerations.

The possible role played by encapsulated and non-encapsulated nerve endings in the skin has been reviewed by Kenshalo and Nafe (16). No one really knows the mechanisms by which sensory receptors convert or transduce the various energies which impinge upon them into nervous impulses. Over the years excision experiments and other types of investigations have produced few answers as to the receptive mechanisms of the skin. Speculations have ranged all the way from free nerve endings to arterio-venous anastomoses, with most attention having been given to end bulbs and corpuscles. The recent works of Quilliam (20), Alvarez-Buylla and Ramirez de Arellana (2), and Lowenstein (17) on the Pacinian corpuscle; of Pfaffman (19), and of Uttal (26) on neural coding, and of Stevens (23) on magnitude estimations lend some credence to the possibility that nerve fibers react with different temporal patterns of neural firing to different types of stimulation, and that the pattern of activity is carried intact to the cortex and there decoded into the various sensations.

D. Some Parameters of Electrical Stimulation

The specific aims of the psychophysical aspects of our research has been to identify the dimensions, parameters, and conditions for pain-free electrical stimulation of the skin commensurate with the demands of an easily learnable coding system. In order to describe the boundary conditions between "pain" and "pulses" we have dealt with problems of intensity, polarity, pulse duration, pulse interval, electrode type, size, and spacing, and with the local skin areas being stimulated. We have been concerned with the possibilities of "skin deafness" and the effects of neurovascular changes on perceptions. The former does not appear to be a problem of concern, the latter is a most important consideration.

The apparatus we have in use was modified from a design developed by Hahn (13) in 1958. Square-wave pulses are generated (from Tektronex equipment) and capacitance coupled to a triode amplifier utilizing a type 24 power triode to stiffen the pulses so that capacitance or resistance of the skin does not distort the waveform. The output of the amplifier is applied to the electrodes at the skin. A meter is connected at the electrodes to keep a constant check of skin resistance. Voltage and current are read on a Dumont 411 type oscilloscope. Data are punched directly on cards and programmed for a computer.

We have made exploratory investigations on some two hundred subjects and extensive studies on a dozen trained observers. Exploration of most of the entire body surface was made for perceptions of pain and painless pulses, under conditions of normal skin temperature and under conditions of vasodilation and vasoconstriction. The only portions of the skin surface not affected by neurovascular changes was found to be the palms of the hands and the soles of the feet. Well defined loci on these areas have a high degree of threshold reliability. Although there are significant differences in sensitivity from one locus to another on a single palm, there is bilateral equivalence in pulse limens for the two palms. Sensitivity to pulses is punctiform regardless of the general skin area stimulated.

We have been able to stimulate sensitive spots on the palms for as long as two hours with no diminution of sensitivity, no change in current or voltage, and without creating local irritations. While no studies on adaptation have been made, we feel that the problems of adaptation are minimal, if important at all in communication. Once a sensitive spot has been stimulated in such a way as to induce pain, such as using very high intensities or very short durations, painless pulses are not possible until adequate time has been allowed for recovery. On the palms of the hands and soles of the feet the "feel" of the pulses seems of uniform quality. Qualitative differences of sensation on the arms, leg, back and other areas are noticeable.

We have made measurements of two point limen of the pulses on various skin areas. This perception is related to positioning of the two electrodes in relation to proximal-distal or medial-lateral direction to each other. For example, the sensation of "twoness" appears to be present on the order of mere millimeters of separation as we go around the arm, but a matter of inches up and down the arm. Because of the critical resistance balance entailed, and of the confusion that results in identifying multiple signals, the simultaneous stimulation of two loci seems to be impractical for coding.

Our original explorations of skin sensitivity to pulses was carried out on the dorsal and ventral surfaces of the forearm. Since we found that a given spot would yield painless pulses on one occasion and pain on another, with the same physical conditions, we considered that some electro-chemical changes might be involved that were related to electrodes. Various electrode materials were tried out, with and without using electrode paste. Carbon electrodes were rejected because of their porous surface and tendency to become contaminated by surface excretions. Tantalum was rejected because of its brittleness and difficulty in machining. Stainless steel was finally chosen from among several inert sources because of its availability, ease of machining, and apparent absence of reactivity. Under the conditions of our work we have found the electrode must have a smooth surface and be of adequate size.

We have used two types of electrodes in our studies. A concentric type was machined from a stainless steel tube with an overall dimension of 33 mm. having a 6 mm. wall thickness. A stainless steel rod was centered inside the tube 6mm. equidistant from the tube walls and held in place by a nylon insulator plug. The other electrode type was the same 6 mm. rod of the concentric electrode where the circuit was completed through a switch to a flat, 3 × 10-inch stainless steel plate on which the subject placed his bare foot. This arrangement allowed the experimenter to switch from one electrode type to the other without the awareness of the subject or the need for changing the position of the electrode.

Using the above equipment Croner (3) has carried out experiments to get measures of the lower threshold of pulse sensation, the upper limit of subjectively determined painless stimulation, and the difference between the two, measured in terms of both voltage and current. Here the experimental design included the independent variables of duration, loci, electrode types, and subjects combined into a four factor factorial design with five repeated measures of voltages recorded at both the lower pulse limen and the upper or pain limen. The same design was repeated using current instead of voltage as the intensity measure. Separate analyses of variance were carried out for each of the limen and the difference between the two.

Pulse durations used in the studies ranged from 1 through 64 milliseconds. Three separate and distinct loci on the palm were used. The electrodes included both the probe and concentric types and twelve subjects were used. Duration, rather than frequency, was used as the primary variable in the determination of both threshold per se as well as the limits within which the stimulation was perceived as painless. Pulses were presented at the rate of one per second. Pilot work had indicated that the range of durations which reliably mediate painless sensation was from around 0.5 msec. to 75.0 msec. Pulse durations shorter than 0.5 msec. result in perceptions of sharpness and pain while durations longer than 75 msec. often lead to unpleasant burning sensations. Throughout the studies, measures of lower threshold of pulse perception, the upper limit of subjectively determined painless stimulation, and the difference between the two were used as the dependent variables.

It was found that both the voltage and current required to reach the pulse limen and the pain limen are highest at the shorter durations. These measures decrease as a logarithmic function of increasing duration. The subjects indicated that the pulses at the shorter durations are qualitatively inferior to those of the longer durations. Under the conditions of these studies it was concluded that a pulse duration of 10 milliseconds would best serve the needs of a communication system. Pulses of this duration are not so long as to significantly increase the time required for communication, and because of their subjective clarity they are easily discriminated.

Our data are in line with some findings from the Virginia Laboratory that pulse frequency must be kept below 8 p.p.s. in order to avoid unpleasant sensations. Frequencies in the 10 to 15 p.p.s. range often brought about feelings of pain, with the likelihood greater for the very short durations. At 20 p.p.s. pain is usually experienced regardless of the length of the pulse duration. Important also is the skin area being stimulated. For example, the same stimulus package that would produce pulses on the forearm induced pain on the calf of the leg. Skin locus is a most important variable in the problem. All of our studies have led us to the conclusion that the palms of the hands and soles of the feet offer the most reliable places to stimulate. Even here it is important to avoid placing the electrodes on a nerve trunk or muscle group. Although there are differences in threshold values obtained from the probe and the concentric electrodes, we favor the latter for convenience.

We now have evidence that painless pulses can be produced quite reliably on the palms of the hands. It remains now to determine the various aspects of the electrical stimulation that can be used in coding. Among the primary dimensions we have intensity, loci, repetition rate, and duration. We have found that even untrained observers can with perfect accuracy discriminate two levels of intensity within a range below the pain level. With practice three levels are easily discriminated. A trained subject has no difficulty in identifying a "high," "medium," and "low" intensity from day to day without resorting to a reference point.

As for loci, it is important to "tune in" those punctiform areas on the palm for the desired sensitivity range. Once this has been done it is possible to reproduce the same pulse perceptions day after day with high reliability. Fortunately, neurovascular changes and perspiration have little effect on these areas after a short interval of time has been allowed for skin resistance to level off.

Repetition rates can be discriminated but there are possible problems in using this dimension in coding. First of all the frequencies exceeding 8-10 p.p.s. feel sharp and uncomfortable and at high intensities painful. Also, an increase in the repetition rate of electrical square-wave stimulation leads to the perception of an increase in intensity as well as frequency, a phenomenon that has been observed with mechanical vibration by Goff (10). Confusion could well result if both intensity and frequency are used as coded dimensions. Also, the use of repeated pulses as a dimension would reduce the speed of communication.

Durations are not nearly as accurately discriminated as are intensities. Trained observers are able to tell changes in subjective quality of the pulses with variations in duration, but the changes are inconsistent and do not correspond with concomitant measured changes in the stimulus. It is possible, however, that two discriminable levels of duration, associated with a secondary dimension of "depth," can be found. Thus, with three levels of intensity, six loci (three on each palm), and two "duration-depth" dimensions, a code made up of 36 discreet symbols seems possible. This would allow for a simple alpha-numeric code of 26 letters and ten digits. More elaborate coding, such as phonetic or the Japanese Kana system, may become possible (1).

E. Problems in Coding

Important to a communication system are questions of stimulus presentations. Whether the stimuli should be presented as discrete elements, regular or irregular, in sequence or not, for a fast response or for a slow response, depends largely on the demands of the output.

The information loop, so common in conventional forms of communication, such as in the completion of a telephone linkage from caller to receiver back to caller, may well have some potentialities in the tactile area. One example is found in the practical problem of the pilot who gets a cutaneous signal on his foot warning him to ease up on the brake pedal. The easing off causes a cessation of the signal and hence the loop has been completed. This is an area for development rather than research as are two kinds of warning situations favorable to tactile stimulation. One involves the "break in" to an on-going activity where attention is demanded above and beyond the activity in which the receiver is engaged. Work in this area should center around the proper signals to use on the proper skin loci. Here one might well wish to compare electro-pulse stimulation with the other sense modalities for, say, speed of reaction. The second type of warning may be described as "drawn out," a prewarning situation as it were. This is exemplified in the case of a person who loses vigilance on the job that does not demand near constant attention. Thus, a warning that attention may soon be demanded can ready the individual for the situation.

The problem of coding and complex information processing becomes more difficult when communication involves the complexities of language. Some of these problems in the design of a cutaneous communication channel are being worked on in our laboratory. They have been described by Gregg (11) as including such problems as rate of reception, code interpretation, memory structure, and cutaneous imagery. Code interpretation requires that discrete segments of the input signal be intermittently sampled by the system. The system must provide a detector mechanism that serves to differentiate among the coded signal elements and that the system operates as a serial machine. A memory structure must contain the associations between signal elements and things for which they stand.

Important new areas are opening in computer simulation of complex human information processing growing out of such work as that of Newell, Shaw, and Simon (18). Conditions affecting the amount of information possible through cutaneous communication is considered of primary importance by Alluisi (1) who says that attention must be directed toward the construction of S-R ensembles. Relating the problems of coding and sensory research Gregg (11) poses such questions as: What is the nature of cutaneous imagery? Given that a number of discrete painless pulses can be produced, what is the size of the immediate memory span for varying numbers of them presented at varying rates? What combinations of levels and dimensions of the cutaneous stimuli provide unique senses "sample points" over variations in rate and intensity? How many distinct symbols based on the cutaneous stimuli can we derive and what is the duration of each? What are the possibilities for simultaneous stimulation over dual cutaneous channels?

Along with investigation of cutaneous phenomena through psychophysical experimentation we believe that information processing concepts, the analysis of language structures relevant to the development of cognitive processes, and the facts of neural coding are basic to the development of cutaneous electro-pulse communication.

F. SUMMARY

This paper deals with some of the psychophysical and coding problems involved in the development of cutaneous electro-pulse communication. We believe that such a system is possible since one can obtain painless pulses under appropriate stimulating conditions suitable for alpha-numeric coding. Attention has been given to the primary dimensions of stimulus duration, intensity, repetition rate, and skin loci.

Our psychophysical findings indicate that it is important to use appropriate electrode types and sizes, to stay within restricted frequency and duration ranges, and to avoid certain skin areas which are affected by neurovascular changes. Measures of voltage and current required to reach the threshold of pulse sensation and the upper limit of painless stimulation decrease as a logarithmic function of increasing duration. Subjects are significantly more variable in reporting pain limen than in reporting pulse limen. Durations around 10 milliseconds require less current to reach threshold and yield larger ranges between pain and pulse limens than do other durations. Pulses of this order of duration give clear-cut sensations and are not so long as to significantly increase the time required for communication using a repetition rate of one pulse per second.

There is bilateral equivalence on selected places of the two palms in terms of threshold values. Under appropriate conditions, it is possible to stimulate the palms for several hours at a time without appreciable diminution of pulse perception and without inducing local irritations. This does not hold for certain other skin areas. Skin deafness probably will not be a problem in working out a practical communication system. Over two hundred subjects we have used in our several studies, and none have failed to perceive the pulse sensations.

Attention is given to coding interums of such problems as rate of reception, code interpretation, memory structure, cutaneous imagery, and complex information processing.

REFERENCES

1. ALLUISI, E. A. On optimizing cutaneous communication: A respectful supplement to some adventures in tactile literacy. *U.S. Army Med. Res. Rep.*, 1960, **424**, 114-130.
2. ALVAREZ-BUYLLA, R., & RAMIREZ DE ARELLANO, J. Local responses in Pacinien corpuscles. *Amer. J. Physiol.*, 1953, **172**, 237-244.
3. CRONER, M. D. Some parameters of cutaneous square-wave electrical stimulation. Unpublished Master's Thesis. Carnegie Institute of Technology, 1961.
4. GAULT, R. H. "Hearing" through the sense of touch and vibration. *J. Franklin Inst.*, 1927, **204**, 329-358.

5. Gault, R. H., & Goodfellow, L. D. An empirical comparison of audition, vision, and touch in the discrimination of temporal patterns and ability to reproduce them. *J. Gen. Psychol.*, 1938, **18**, 41-47.
6. Geldard, F. A. Adventures in tactile literacy. *Amer. Psychol.*, 1957, **12**, 115-124.
7. ———. Some neglected possibilities of communication. *Science*, 1960, **131**, 1583-1588.
8. Gibson, R. G. Electrical stimulation of the skin—on purpose. Paper read at Virginia Academy of Science, May 1960.
9. Gilmer, B. v. H., & Haythorne, S. R. Cutaneous pressure-vibration spots and their underlying tissues. *Arch. Neurol. & Psychiat.*, 1941, **46**, 621-648.
10. Goff, G. D. Differential discrimination of frequency of cutaneous mechanical vibration. Unpublished Doctoral Dissertation, University of Virginia, 1959.
11. Gregg, L. W. Some coding problems in the design of a cutaneous communications channel. *U.S. Army Med. Res. Rep.*, 1960, **424**, 85-102.
12. ———. Engineering psychology, in *Industrial Psychology*, ch. 16 (Gilmer, B. v. H., Ed.), New York: McGraw-Hill, 1961.
13. Hahn, J. F. Cutaneous vibratory thresholds for square-wave electrical pulses. *Science*, 1958, **127**, 879-880.
14. Hawkes, G. R. (Ed.) Symposium on cutaneous sensitivity. *U.S. Army Med. Res. Rep.*, 1960, **424**, 1-165.
15. Hawkes, G. R., & Warm, J. S. ΔT for electrical cutaneous stimulation. *J. Psychol.*, 1961, **51**, 263-271.
16. Kenshalo, D. R., & Nafe, J. P. Receptive capacities of the skin. *U.S. Army Med. Res. Rep.*, 1960, **424**, 1-25.
17. Lowenstein, W. R. Biological transducers. *Scientific Amer.*, 1960, **203**(2), 98-108.
18. Newell, A., Shaw, J. C., & Simon, H. A. The elements of a theory of human problem-solving. *Psychol. Rev.*, 1958, **65**, 151-166.
19. Pfaffman, C. The afferent code for sensory quality. *Amer. Psychol.*, 1959, **14**, 226-232.
20. Quilliam, T. A. Dimensions of the ultra structure of the Pacinian corpuscle. *J. Physiol.*, 1957, **137**, 2-3.
21. Sigel, H. Prick threshold stimulation with square-wave current: A new measure of skin sensibility. *Yale J. Biol. Med.*, 1953, **26**, 145-154.
22. Smith, J. M. Which hand is the eye for the blind? *Genet. Psychol. Monogr.*, 1929, **5**, 213-222.
23. Stevens, S. S. Measurement and man. *Science*, 1958, **127**, 383-389.
24. von Békésy, G. Similarities between hearing and skin sensations. *Psychol. Rev.*, 1959, **66**, 1-22.
25. von Frey, M., & Kiesow, F. Ueber die function der tastkorperchen. *Z. Psychol.*, 1899, **20**, 126-163.
26. Uttal, W. R. A comparison of neural and psychophysical responses in the somesthetic system. *J. Comp. Physiol. Psychol.*, 1959, **52**, 485-490.
27. Weiner, N., Wiesner, J. B., David, E. E., Jr., & Levine, L. Operation "Felix." *Quart. Progr. Rep.*, Res. Lab. of Electronics, Mass. Inst. of Tech., 1949-51.

Department of Psychology
Carnegie Institute of Technology
Schenley Park
Pittsburgh 13, Pennsylvania

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 223-229.

INTRACTABLE ASTHMA: CONFLICT OF PERIOD OF TOILET TRAINING*[1]

Biological Laboratory, Cold Spring Harbor, New York and Community Hospital, Glen Cove, New York

---

HAROLD A. ABRAMSON

There are many chemical, physical, infectious, and allergic reactions which produce expiratory wheezing in the chests of children and adults. The physical sign of wheezing may be very mild and just distinguishable, or it may be accompanied by respiratory distress characterized by a rather forceful expiration of trapped air. In general, this respiratory syndrome is called asthma, irrespective of its etiology and severity. Some asthmatic symptoms may be accounted for quite simply. After inhalation of chlorine gas, for example, typical asthmatic breathing ensues. Similarly, after running uphill on an extremely cold day, I personally had one single attack of asthma. Often upper respiratory infections with low fevers produce a type of "infectious" asthma. Certainly, typical allergic responses to inhalation (or injection) of pollens and danders, as well as to ingestion of many foods, are allergic asthma of the classical and common variety. These usually respond to elimination of an allergen or to other forms of therapy. Previous psychodynamic formulations are reviewed elsewhere (8), (9).

Even when physical, etiological factors of the foregoing types stand in clear relationship to the presence of asthma and even though the asthmatic symptoms respond to therapy by drugs, sedation, and immunological procedures, emotional factors may also be present. How important these emotional factors may be must be evaluated not only in each case or at any given time, but also at different periods in the patient's life and actually from day to day, hour to hour, moment to moment.

The question which really confronts us is, therefore, not whether emotional factors are involved, but how important these emotional factors are in any case of intractable asthma. When a child regularly gets asthma during the pollen season and the asthma responds to conventional treatments, it it reasonable to assume that the child's response is primarily immunologic

---

* Received in the Editorial Office on April 26, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation has been supported in part by a grant from the Foundation for Research in Pulmonary Disease, New York City.

in origin. In the simplest cases, however, psychological factors may be significant. For example, in the case of the single asthmatic attack produced in myself by violent exercise during extremely cold weather, the asthmatic response, although induced by cold, was also controlled by the nervous system. The anxiety which I felt on being immobilized by the asthmatic attack could have led to more serious consequences. As a matter of fact, I remained in the same spot for about ten minutes, thinking matters over, patiently waiting for the bronchial spasm to disappear. Perhaps if I had not understood the mechanisms involved in asthma, subsequent asthmatic attacks, under similar or other circumstances, might have occurred.

Let us now discuss the symptoms of asthma in those individuals in whom immunological factors cannot be clearly utilized in therapy and especially in those patients who do not respond to treatment, which Peshkin (6) notably has termed intractable asthma. We must explore the nature of the psychological vectors and make the assumption that further search for allergens significant in a program of therapy may be fruitless as compared with the search for emotional vectors which have changed the patient so that the patient no longer responds to ordinary treatment procedures. The careful research of Peshkin (4) has proved, without doubt, that a large number of randomly chosen cases of intractable asthma in children (intractability being defined as hitherto: no longer responding to conventional treatment), lose all or practically all their asthma within a very short period after being removed to an institution and being separated temporarily from their parents. Peshkin has called this separation "parentectomy." The almost miraculous recovery of about 40 per cent of the children removed from their homes emphasizes that, in spite of the intractability of the asthma, no important morphological and pathological process can be present and that these children who recover rapidly must have had an asthmatic syndrome whose governor must be maintained in emotional areas of the personality rather than in specific immunological areas of hypersensitivity. This communication deals with one group of seriously ill asthmatic children who do not respond to conventional methods of therapy.

The intractability of asthma in both children and adults has led me to search for some general psychological process which would be consistent with the fact that psychoanalysis has not as yet provided us with a satisfactory solution to the problem of intractable asthma. Although psychoanalysis provides us with the raw material for some psychodynamic processes, conscious attitudes also supply us with important data. For example, children make subtle substitutions of a more palatable behavior pattern for one unpalatable to the parent. Sucking the edge of the pillow case at bedtime is more acceptable than thumb sucking; rubbing the fur of a toy animal is less threatening to the parent than self-stimulation by the child of one of its erotic zones. Indeed, the itching, disturbed, eczematous child is often quieted when he scratches his mother who rarely objects. Undoubtedly, one of the most trying periods of childhood development for both parents and children is the period of toilet training. Consciously and unconsciously, the child will try to control the parent's displeasures and even utilize his own illness for this purpose. Behavior disorders in this particular area of development are well known. Counteridentification processes within the parents are particularly apt to be activated, and unconscious disapproval by the parents of the child during this period is often uncontrollable (2).

English and Pearson (3) point out that when the parent shows displeasure at the passage of gas through the lower end of the intestinal tract, the child may, on relinquishing this pleasure before he is ready to do so, employ the upper end of the intestinal tract for the production of obscene and scatological speech. Thus, nursery school children have epidemics of obscenity, which, although shocking adults, is more acceptable than the public passage of flatus. A child may also become apprehensive lest he yell or scream in a public place. More important to the psychodynamic formulation to be proposed here is the effect of too early a relinquishment of anal pleasure on the respiratory pattern. A threatening parent may force the child during toilet training to regress to the infantile method of breathing. The respirations are not regular and catching the breath is common. This leads to difficulty in speech and stuttering. According to this psychodynamic formulation, stuttering may represent a conflict between a desire to express hostility toward the parent by the passage of gas from the body and the desire to please the parent by not performing an improper action. In this way the conflict is shifted from one end of the intestinal tract to the other end. The internal gases play a role now in the function of speech, and the conflict is shifted, therefore, to an expression by the respiratory function.

Let us examine what a child does when the stress and threat of a parental command imposes on him the task of desperately trying to pass a stool. The child may take a deep breath, strain the rectoanal sphincter, get red in the face, and even turn blue trying to push out a fecal mass that may or may not be ready for passage. It has not been uncommon in my practice for parents to confess to me that if their children did not move their bowels on command, so to speak, their children (16 months and upward) were struck with the hand, a ruler, a strap, or were otherwise threatened. Consider what might happen in the allergic child with his special immunological makeup, when he, both consciously and unconsciously, has learned that the wheezing of asthma, another type of forcing out of air, may reduce the threatening nature of the relationship due to parental displeasure resulting from the child's failure to pass a fecal mass and comply with the parent's needs. This failure to comply and force out a stool may now lead the child through a motion similar to moving the bowels—lowering the diaphragm, tensing the abdomen and chest muscles, and forcing air out from the other (proximal) end of the intestinal tract. It is remarkable that during many attacks of asthma individuals will purse the lips in the prolonged expiration in the same way that the anal sphincter is pursed during the passage of the fecal mass. Perhaps, unconsciously, the passage of flatus through the lower end of the intestinal tract signals the child that the passage of air from the upper end of the intestinal tract during the asthmatic attack might provide a device for managing a distorted relationship with the threatening parent. That forcing air out in an asthmatic wheeze controls the threatening parent validates for the child the value of the unconscious signal induced by passage of flatus. Purcell (7) has studied the records of the Jewish National Home for Asthmatic Children at Denver covering the 1957-58 discharge files. He compared 14 children who rapidly lost their symptoms of asthma on being admitted to the Home with another 14 children who did not lose their asthma as readily and were dependent on corticosteroid therapy. The group of children who lost their asthma rapidly had a mean age of onset of about 39 months, whereas those children who were dependent on steroid therapy for relief had a mean age of onset of 26 months. In other words, the more intractably asthmatic the child, the closer was the age of onset with the period of toilet training. I recognize that asthma may occur in infancy before toilet training begins. This group of infantile asthma does not fit into this phase of intractable asthma. For example, in Figure 1 are plotted data obtained from Peshkin's classical paper on intractable asthma. Note in the Figure that the frequency of onset of intractable asthma is highest during and immediately before and after the period of toilet training. The frequency of onset then sharply decreases and essentially becomes zero with puberty. In other words, the further away the age of onset of childhood asthma from the period of toilet training, the greater the probability that it will be treatable by conventional methods.

Another way in which the intestinal tract is connected with asthma is the well known fact that administration of ipecac and the subsequent vomiting relieves asthma in many children. This may be symbolic passage of a fecal mass. Psychiatrists have long wondered why aminophllyn suppositories are often successful when drugs equally active pharmacologically failed to relieve asthma when administered by routes other than the anal one. This problem remains unsolved.

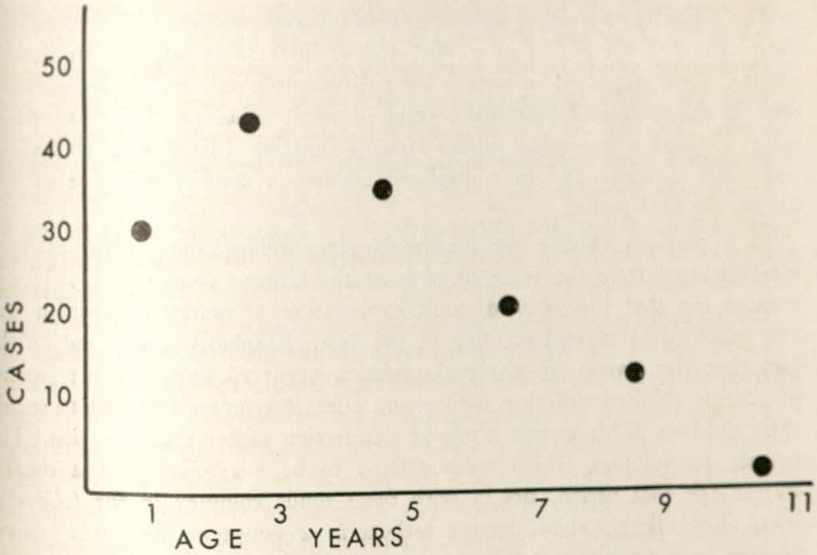


FIGURE 1
Data of Peshkin. The number of cases of intractable asthma admitted to the Jewish National Home for Asthmatic Children is plotted against the average time of onset of asthma. Note that the period of toilet training is the time when the number of cases is at a peak. After this period the age of onset diminishes linearly with time, almost reaching zero by puberty.

An old slogan amongst chest men is, "A barium enema often clears the lungs." It would be of interest to treat a group of intractably asthmatic children with a series of barium enemas or other types of enemas which would enable the child to know that periodically, under medical supervision and understanding, a mass could be forced through the anal sphincter.

Dr. Peshkin (5) has informed me that a significant number of children are violently opposed to any medication inserted rectally. Medication does not work under these circumstances.

In the last few weeks a patient said to me, "I am getting some asthma, and my stomach is contracting, too." In fact, the pre-asthmatic symptoms are often connected with gastrointestinal spasm.

The difficulty of treating asthma by conventional psychotherapeutic methods correlates with the period of conflict and of intractable asthma. Intractable asthma, then, would fit in with a compulsive autonomic mechanism so often seen in the intractably asthmatic child or adult who, unrelieved by drugs, solves through a somewhat similar striated function (diaphragm, abdominal, and chest muscles) and in a displaced fashion, a conflict with the parents during the anal period.

In a previous volume the mother-daughter relationship in the allergies was discussed from the point of view of the Cronus complex (1). It was pointed out that the psychodynamic explanation of maternal rejection was not validated by careful scrutiny of the data. Emphasis was placed on the fact that the reverse of maternal rejection occurred and that the parents of allergic children suffering with severe allergic symptoms tended to engulf their children. This process is one of introjection rather than rejection. The present formulation, which holds asthma to be a resolution of a conflict during the anal period, fits in with the Cronus complex in the following way. The allergic child, having asthma as a symptom, mobilizes anxiety because he is unable to fulfill the demands of the introjecting parent. The child, unconsciously, therefore, knows that the parent will be pleased under certain circumstances by the passage of gas and other matter from the distal end of the intestinal tract and tests the parent with the allergic symptom of asthma. Asthma is readily produced by older children consciously. The threat to the child by the introjecting parent is replaced by the production by the child of anxiety in the introjecting parent. The child, suffering from anxiety because of the need to fulfill conflicting parental goals, retaliates by producing anxiety in the parent.

SUMMARY

During the period of toilet training the child must mobilize a highly coordinated task of moving the bowels on parental command. The physiological process of lowering the diaphragm, tensing the abdomen and chest muscles corresponds, in general, to the process of forcing air through the proximal end of the intestinal tract during an asthmatic attack. It is proposed that the intractable asthma of children is a device used by the child to transform the threat of the parent's displeasure during toilet training to

one of anxiety production in the parent by utilizing the proximal rather than the distal end of the intestinal tract. Thus, the compulsive, intractable asthma of the child resolves certain conflicts imposed by the parent during the period of toilet training. Data supporting the psychodynamic formulation are presented.

REFERENCES

1. ABRAMSON, H. A. The Patient Speaks. Vantage Press, New York, N.Y. 1956
2. ———. Lysergic Acid Diethylamide (LSD-25): XXXII. Resolution of Counter-Identification Conflict of Father during Oedipal Phase of Son. *J. Psychol.*, 1961, **51**, 33-87.
3. ENGLISH, O. S., & PEARSON, G. H. J. Emotional Problems of Living. W. W. Norton & Co., Inc., New York, N.Y. 1945.
4. PESHKIN, M. M. Intractable Asthma of Childhood. *Int. Arch. Allergy,* 1959, **15**, 91-112.
5. ———. Personal Communication.
6. PESHKIN, M. M., & ABRAMSON, H. A. First National Seminar of Regional Medical Consultants. *Ann. Allergy,* 1958, **16**, 473-509.
7. PURCELL, K. Second National Seminar of Regional Medical Consultants. *J. Children's Asthma Res. Inst. Hosp.,* 1961, **1**, (In press.)
8. ROSS, N. Somatic and Psychiatric Treatment of Asthma. (H. A. Abramson, Ed.), Williams and Wilkins, Baltimore, Md. 1951. p. 655.
9. SELINSKY, H. Emotional Factors Relating to Perennial Allergy. *Ann. Allergy,* 1960, **18**, 886.

133 E. 58th Street
New York 22, New York

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 231-239.

THE INFLUENCE OF A DISTORTED GROUP NORM UPON JUDGMENTS OF ADULTS AND CHILDREN*[1]

Department of Psychology, University of California, Berkeley, California

READ D. TUDDENHAM

In the years since Asch first reported on the behavior of subjects faced with an apparently erroneous majority, many aspects of the yielding situation have been elucidated by Applezweig (1), Bass (3), Blake (5), Crutchfield (6), Deutsch (7), Tuddenham (11, 12, 13, 15, 16) and others. These findings constitute an unusually coherent body of fact,[2] but much of this work has been limited to college students. In those instances where other groups have been used, there have occurred certain inconsistencies in regard to sex differences. Thus, Crutchfield discovered and the present writer independently confirmed that college women are more susceptible to the false norm than college men. However, Crutchfield found an unexpected degree of independence in a group of older women, and the writer found that with a group of subjects in their thirties, sex differences in visual discrimination were small, and in other types of content, virtually non-existent.

Although studies of children should be especially enlightening with regard to the origin of sex differences, only Berenda (4) has investigated children, using an experimental situation patterned closely on Asch's procedure in which yielding and independence are discrete categories of response. She reports children yielded more frequently than college students, and younger children more frequently than older ones. However, "complete inde-

---

* Received in the Editorial Office on May 8, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This paper is an abridgment of Technical Report No. 11, ONR Project NR 170-159 (Tuddenham, 1960). It is one of a series which demonstrate that modification of judgment toward a group norm is complexly determined by the nature and clarity of the task, the composition of the group and the qualities S attributes to them, the magnitude of the distortion, and such basic variables of S's own personality as his value system, his intelligence and his self confidence. The position has been taken that the psychological processes underlying yielding and independence are fundamentally alike, though following from different weightings of the relevant variables, and that no one wholly disregards either the physical reality or the conflicting social consensus. The financial support of the Office of Naval Research is gratefully acknowledged. However, the Navy Department is not responsible for the author's findings or conclusions.

[2] See the recent report of Bass (3) for a summary and theoretical integration.

pendence" and "complete yielding" occurred in all groups. She does not report upon sex differences.

The present study is concerned with the yielding phenomenon in children of about the same age as Berenda's, in comparison with college students. However, the writer's experimental procedure permits a continuous range of responses on the individual item under conditions of strict control and comparability from group to group, and is hence somewhat better adapted to measuring *degrees* of distortion of response between the extremes of yielding and independence.

The primary purposes of this investigation are to assess the relative susceptibility to yielding of children and adults, and to determine whether the sex difference usually reported obtains also among these younger subjects. A secondary purpose is to develop and calibrate a stimulus series suitable within rather broad limits for making age and cross-cultural comparisons. For this reason, the opinion and information items contained in Standard Experimental Series No. 1, and used in the previously reported experiments, were dropped. A new standard stimulus series was constructed wholly out of visual discrimination items, and the verbal formulae simplified in phrasing and vocabulary. While the final series (SES-2) is shorter over all, it contains a longer and more varied set of visual judgments than before.[3]

A. METHOD

1. *Subjects*

Subjects in the college group were members of lower division psychology and ROTC courses at the University of California, drawn in the manner described in the earlier experiments. Thirty-two men and 33 women were given the experimental treatment. Thirty men and 31 women were controls, i.e., reported their judgments independently and without knowledge of any norm.

Subjects in the children's group were members of a six-week summer recreation school program conducted on Campus.[4] Children attending the school were drawn from middle-class families able to pay the tuition fee, —largely business and professional people and members of the University staff. The experimental group consisted of 35 boys and 35 girls. A children's control group of 24 boys and 25 girls was treated exactly the same as the college controls. The youngest child in either group was nine years, nine months old, the oldest had not reached his thirteenth birthday. (Boys' mean = 11 years, 4.9 months; girls' mean = 11 years, 3.4 months.)

---

[3] A complete description of this series is contained in Tuddenham (14).

[4] Grateful acknowledgement is made to Mr. Heber Newsom and to the staff of the Recreation School for permission to test these children.

Subjects were run in groups of the same sex and approximate age. The college subjects were ordinarily unacquainted, in contrast to the children, but in all groups, subjects were routinely introduced to each other. Experimental and control groups were run concurrently, as described in previous studies.

2. *Apparatus*

The apparatus has been described previously (Tuddenham, 11). Five subjects sat next to each other, each at a panel screened from his neighbors. Nine-part multiple-choice problems were projected on the wall opposite, and the subject, under the reiterated injunction to be accurate, responded to each by pressing that one in a row of nine toggle switches on the panel which corresponded to his choice. Under control conditions, subjects responded in unison with no knowledge of each other's choices. Under experimental conditions, subjects responded in turn (the order changing from trial to trial) and signals lights on the panels ostensibly showing the choices of the others but in fact faked by the experimenter, were provided "to tell the subject when to take his turn." On non-critical trials, the faked choices were reasonable ones. On critical trials (every third slide) the faked choices lay outside the range of responses given during preliminary standardization of the various items.

3. *Items*

All items involved visual perception but were as varied as possible in content, including comparisons of line length, area, slope, convergence, trajectory, proportion, etc. The series as a whole consisted of 4 trials for demonstration, followed by a series of 46 trials, the 15 critical items being spaced evenly through the series. One slide used for a critical trial failed to yield the necessary consensus among control group subjects and was dropped from all analyses.

4. *Scoring*

As in previous studies, each subject's raw score was simply the sum of his switch number choices on the 14 critical trials included. For certain analyses, these raw scores were converted to standard scores, based on the mean (46.5) and *SD* (2.1) of the college control group. Inasmuch as there was not a significant difference between them in mean or variance, men's and women's control data were combined to calculate the statistics required for the z-score transformation. The degree of yielding was expressed also in quasi-absolute units as a percentage, taking the control group median

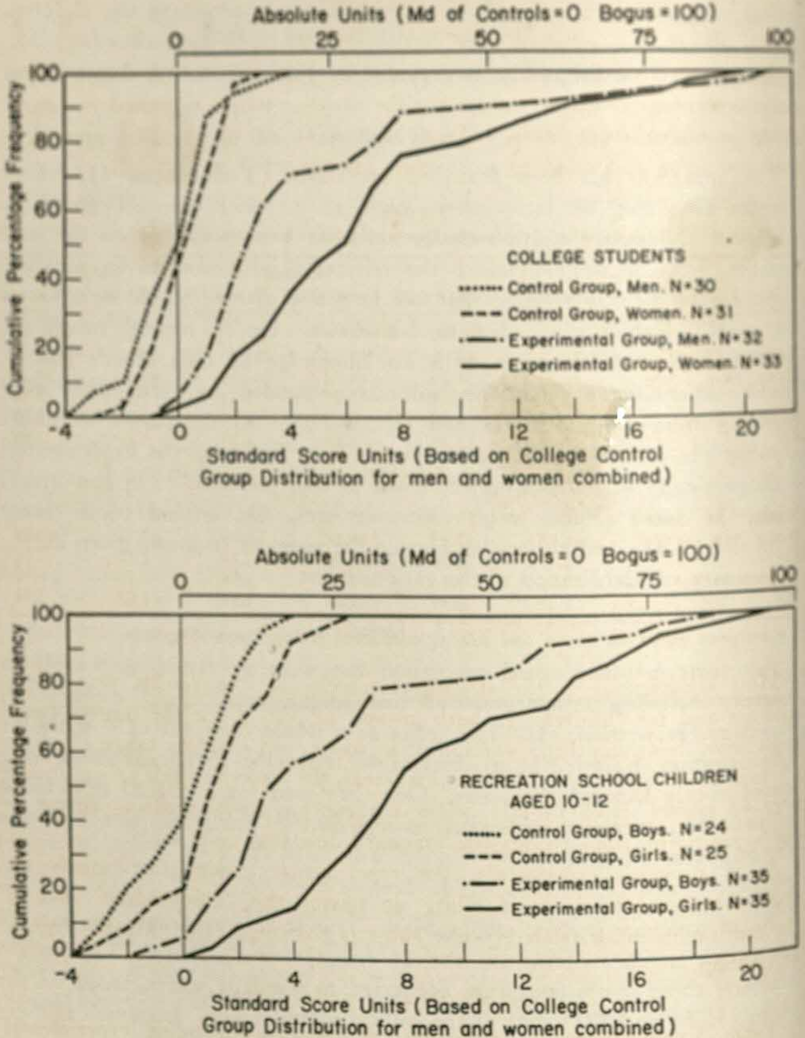


FIGURE 1

CUMULATIVE DISTRIBUTIONS OF YIELDING SCORES IN VARIOUS GROUPS as a zero point, and the score corresponding to consistent agreement with the distorted norm as 100.

B. Results

1. *Score Distributions*

An overview of the results is provided by Figure 1, which shows cumulative percentage frequency curves for the various groups, expressed in control group standard score units. Means and *SD*'s of the various groups in both raw score and standard score units are presented in Table 1.

TABLE 1
Means and Standard Deviations of Various Groups

	Men		Women		Combined	
	Control	Exper.	Control	Exper.	Control	Exper.
			College Students			
Number of cases	30	32	31	33	61	65
Mean raw score	46.23	56.25	46.32	60.14	46.50	58.22
SD	2.30	7.93	1.87	7.51	2.10	7.74
Mean *z*-score[a]	—0.12	4.64	0.15	6.50	0.00	5.58
SD	1.01	3.78	0.89	3.58	1.00	3.69
			Children			
	Boys		Girls		Combined	
Number of cases	24	35	25	35	49	70
Mean raw score	46.75	57.13	49.14	64.94	47.97	60.99
SD	2.59	7.54	2.70	7.74	2.78	8.10
Mean *z*-score[a]	0.12	5.06	1.26	8.78	0.70	6.90
SD	1.23	3.59	1.29	3.69	1.32	3.86

[a] *z*-scores based on mean and *SD* of combined college control group.

Most conspicuous is the striking similarity of the results for college students and for children. In both groups, subjects given the control treatment show relatively little variability in scores—an extreme range of about 15 raw score points in 14 items, whereas the experimental groups range over 45 points, from slightly below the control group medians almost to the bogus norm, some twenty-odd standard deviation units away.

As in previous experiments, the experimental group distributions are decidedly skew, most scores piling up nearer the "independent" end of the continuum, with relatively few subjects showing really extreme degrees of yielding.

2. *Experimental vs. Control Groups*

Table 2 summarizes the differences between corresponding experimental and control groups, together with critical ratio tests of the differences in means, and *F* tests of the differences in variances. In every instance, the experimental group has a very significantly greater mean score and a very significantly greater degree of score variability than does the corresponding control group. Clearly, the group pressure introduced in the experimental treatment has an important influence upon the behavior of all groups, regardless of age or sex. This effect, here based on summation across 14 items, is still marked when individual items are considered. Forty of the 56 such experimental-control comparisons yielded *t*s greater that 2.5.

TABLE 2
DIFFERENCES IN YIELDING BETWEEN CONTROL AND EXPERIMENTAL GROUPS

	College Students				Children			
	Men		Women		Boys		Girls	
	Control	Exper.	Control	Exper.	Control	Exper.	Control	Exper.
N	30	32	31	33	24	35	25	35
Mean	46.23	56.25	46.82	60.14	46.75	57.13	49.14	64.94
CR	6.85*		9.87*		7.52*		11.16*	
SD	2.30	7.93	1.87	7.51	2.59	7.54	2.70	7.74
F	11.86*		16.14*		8.47*		8.19*	

* Significant at .01 level.

3. *College Students vs. Children*

Table 3 summarizes the age differences observed. Differences in mean score between the college men and the boys are small and non-significant in both experimental and control groups. Differences in mean score between the college women and the girls reach the 1 per cent level in both experimental and control groups, the girls being less accurate than the older women. These differences are not large in the absolute sense,—a cumulative difference over 14 trials of 2+ raw score points in the control group, and 4+ points in the experimental. Age differences in variability approach significance only among female controls.

TABLE 3
DIFFERENCES IN YIELDING BETWEEN COLLEGE STUDENTS AND CHILDREN

	Control Groups				Experimental Groups			
	Males		Females		Males		Females	
	Coll.	Children	Coll.	Children	Coll.	Children	Coll.	Children
N	30	24	31	25	32	35	33	35
Mean	46.23	46.75	46.82	49.14	56.25	57.13	60.14	64.94
t	0.77		3.64**		0.46		2.60**	
SD	2.30	2.59	1.87	2.70	7.93	7.54	7.51	7.74
F	1.27		2.09*		1.11		1.06	

* Significant at .05 level.
** Significant at .01 level.

Within the restricted range of the children's experimental groups, age is unrelated to yielding, the correlations being .16 for girls and —.14 for boys.

Table 4 presents the findings on sex differences. Among adult experimental groups, the difference in means produces a *t* of 2.03, close to the 2 per cent level (one tail), the women yielding more than the men. When no pressure is present, the women are virtually as accurate (*t* for control groups = 1.10). Among children, the girls are significantly less accurate (control *t* = 3.16) and significantly more susceptible to the group pressure (exp. *t* = 4.28) than are the boys.

TABLE 4
DIFFERENCES IN YIELDING BETWEEN MALE AND FEMALE SUBJECTS

	College Students				Children			
	Control		Experimental		Control		Experimental	
	Men	Women	Men	Women	Boys	Girls	Boys	Girls
N	30	31	32	33	24	25	35	35
Mean	46.23	46.82	56.25	60.14	46.75	49.14	57.13	64.94
t	1.10		2.03*		3.16**		4.28**	
SD	2.30	1.87	7.93	7.51	2.59	2.70	7.54	7.74
F	1.51		1.11		1.09		1.05	

* Significant at .05 level.
** Significant at .01 level.

Differences in variability between comparable sex groups are small and none are significant.

C. Discussion *

As in previous reports, the major source of score differences is the experimental treatment, which produces much larger means and variances than in comparable control groups, regardless of age or sex. These results are in good agreement with those presented previously (see Tuddenham, 11, Table 2), for a visual series of ten items.

Age differences in scores do not loom very large. The boys, college men and college women behave in very similar ways, both with respect to means and *SD*s. Indeed, the boys' means fall between those for men and women in both experimental and control groups. The school girl controls are significantly less accurate than the college women, but this difference is not very impressive in absolute terms. Among experimental subjects, the girls have significantly higher yielding scores than do the college women, but owing to greater variability in experimental groups, the *t* is actually somewhat smaller than among controls. The data thus agree with Berenda's conclusion that children yield more than adults, but the magnitude of the difference is less than she reports. Quite possibly her face-to-face confrontation of the subject by an opposing majority is enough more threatening than our milder situation to account for the difference between her results and ours.

Sex differences are consistent, the female subjects earning significantly larger mean yielding scores than do the males. This difference is *more* conspicuous among the children than among the adults. The girl controls are also somewhat less accurate than the boys, who resemble the adult controls. The outcome of this study is thus in agreement with the great bulk of studies which find women more susceptible to social influence than men. However, it was not anticipated that this sex difference would prove to be relatively greater than the age difference between 10- and 11-year-olds and college students drawn from the same general social level. There is no reason to assume that the sex difference is biologically based, but the number of years of differential training of boys and girls necessary to produce it is unknown. It would be useful to know whether or not the difference obtains among really young children, and to have other information upon the subjects, their intelligence, self-confidence, sociability, etc., which was not available to us. Unfortunately, our technique, which is abstract and relatively impersonal, and which requires children to answer "strictly in turn," is not well adapted to investigating yielding behavior among subjects much younger and less docile than those studied here. Cross-cultural studies of groups among whom sex-typing proceeds along different lines remain, however, a feasible approach to investigating the intricate relationships among cultural sex-stereotype, individual personality, and social behavior of this kind.

D. SUMMARY

Groups of college students and of 10- to 12-year-old children of similar background were tested for disposition to yield to a distorted group norm which conflicted with the physically "right-appearing" answer in a series of visual discrimination problems. Under control conditions, the girls were somewhat less accurate than the boys and the college subjects, both men and women. Under experimental conditions, all groups showed significantly greater yielding score means and variabilities than did controls. Children yielded more than the adults, but the difference between age groups was less impressive than the sex difference. The causes of the observed sex differences are hypothesized to lie in the differential training given boys and girls in

our society. Investigations in other cultures which have different sex stereotypes might be expected to throw further light on this important problem.

REFERENCES

1. APPLEZWEIG, M. H., & MOELLER, G. Conforming behavior and personality variables. Technical Report No. 8, Contract NONR 996 (02), Connecticut College, New London, 1958.
2. ASCH, S. E. *Social psychology.* New York: Prentice-Hall, 1952.
3. BASS, B. M. Conformity, deviation, and a general theory of interpersonal behavior. Technical Report No. 20, Contract N7ONR 35609, Louisiana State University, Baton Rouge, 1960.
4. BERENDA, R. W. The influence of the group on the judgments of children. New York: King's Crown Press, 1950.
5. BLAKE, R. R., HELSON, H., & MOUTON, J. S. The generality of conformity behavior as a function of factual anchorage, difficulty of task, and amount of social pressure. *J. Pers.*, 1957, **25**, 294-305.
6. CRUTCHFIELD, R. S. Conformity and character. *Amer. Psychologist,* 1955, **10**, 191-198.
7. DEUTSCH, M., & GERARD, H. B. A study of normative and informational social influences upon individual judgment. *J. Abnorm. Soc. Psychol.,* 1955, **51**, 629-636.
8. JONES, E. E., WELLS, H. H., & TORREY, R. Some effects of feedback from the experimenter on conformity behavior. *J. Abnorm. Soc. Psychol.,* 1958, **57**, 207-213.
9. TUDDENHAM, R. D., MACBRIDE, P., & ZAHN, V. Studies in conformity and yielding I. Development of standard experimental series—1. Technical Report No. 1, Contract NR 170-159, University of California, Berkeley, 1956.
10. TUDDENHAM, R. D. The influence upon judgment of a grossly distorted norm. Technical Report No. 2, Contract NR 170-159, University of California, Berkeley, 1957.
11. ———. The influence of a distorted group norm upon individual judgment. *J. Psychol.,* 1958, **46**, 227-241.
12. ———. The influence of an avowedly distorted norm upon individual judgment. *J. Psychol.,* 1958, **46**, 329-338.
13. ———. Correlates of yielding to a distorted group norm. *J. Pers.,* 1959, **27**, 272-284.
14. ———. The influence of a distorted norm upon judgments of children and adults. Technical Report No. 11, Contract NR 170-159, University of California, Berkeley, 1960.
15. TUDDENHAM, R. D., & MACBRIDE, P. D. The yielding experiment from the subject's point of view. *J. Pers.,* 1959, **27**, 259-271.
16. TUDDENHAM, R. D., MACBRIDE, P., & ZAHN, V. The influence of the sex composition of the group upon yielding to a distorted norm. *J. Psychol.,* 1958, **46**, 243-251.

Department of Psychology
University of California
Berkeley 4, California.

Published as a separate and in *The Journal of Psychology*, 1961, 52, 241-250.

MEASUREMENTS OF SIMULTANEOUS BRIGHTNESS CONTRAST ACROSS THE RETINA*[1]

Department of Psychology, Michigan State University

WILLIAM R. MACKAVEY, S. HOWARD BARTLEY, AND CARMINE CASELLA

A. PROBLEM

Brightness can be altered by several means not involving an intensive change in the energistic impingement on the eye. For example, a change in the location of the image on the retina or in the visual angle subtended may be sufficient to alter brightness. Studies of simultaneous brightness contrast (6, 7, 9, 17) have shown that initial brightness can also be altered by placing a second more intense target in the field nearby. Although the retina has for some time been regarded as prominently involved in the production of such effects (18) it has only been recently that basic physiological evidence has been provided on this matter. Hartline, Wagner, and Ratliff (12) have shown that the rate of discharge from an illuminated ommatidium, the receptor unit of the compound eye of *Limulus polyphemus*, can be markedly reduced by illuminating a second nearby ommatidium. The fibers which form a horizontal plexus behind the ommatidia have been suggested as the mediating structures inasmuch as their severance irreversibly abolishes this capacity for inhibitory action. A straightforward adoption of this demonstration would suggest that it is the horizontal cells of the human retina which mediate simultaneous brightness contrast.

Unquestionably, data from more than one type of psychophysical situation have relevance in a discussion of the mechanisms involved in simultaneous brightness contrast, e.g., color contrast (10), formation and obliteration of seen edges (4, 13, 19, 20, 21), and temporal features in contrast (1, 2, 3). In fact, there is a tendency on the part of some, perhaps unwittingly, to treat certain of the observations made in these situations not only as related but as equivalent. The reduction in the *brightness* of a target which characterizes metacontrast, for example, is not always distinguished from the reduction in *visibility* which occurs in studies of oblit-

---

* Received in the Editorial Office on May 9, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The present study was done under a National Institutes of Health Postdoctoral Fellowship (MF-10, 138-C1) to the senior author. Likewise, the costs of publication were defrayed by a grant (NIH 71-2207) from the same auspices.

eration of edges. Our immediate concern, however, will be with a single variable within a single set of conditions and its implication for an explanation of contrast phenomena, viz., the relation of using one or both eyes to the effects produced. It is presumably possible to reduce the brightness of a target by presenting a second target, appropriately timed and positioned, to the opposite eye. This is to admit a functional relation between the two sides of the visual system at some level beyond the chiasm, but it is not necessarily to demonstrate specifically what this relation may mean in various cases. Some visual functions are what they are by reason of both hemispheres being involved. Some functions do not differ whether one or the other is involved and some functions may possibly be dependent on a single hemisphere. The inclusion of the second eye in some experimental task situations is sometimes only for removing certain unwanted influences between two portions of the same retina, such as veiling stray retinal illumination, or neural interactions.

The question is not so much whether a given function such as we are studying here is "produced" in one part of the visual system or another, as it is which portions of the system when differently dealt with alter the outcome (response) in question in a detectable and orderly way. For example, the retina constitutes a spatial spread or carpet all parts of which may be impinged upon uniformly, or some parts of which may be impinged upon differently than others, or only restricted portions of which may be involved at all. Naturally there is an anatomical and a conductional relation between the various portions of the retina and the cortex, etc. These connections are of such sorts that only certain experimental manipulations of combination and exclusion can be made. While it is the wish of investigators to be able to ascribe to certain portions of the visual system certain processes and functions, this may not always be possible.

In the present case, one is not able to manipulate the location of the impingement upon restricted portions of the retina in such ways as to relate the retinal effect to one or the other of the two cerebral hemispheres at will. All that one can determine directly in this way is whether or not the outcome being studied is varied by such a procedure. If it is, there is the customary tendency to attribute the seat of the function to the brain. If the outcome remains unaltered, the function is attributed to the retina. This is a tacit way of denying that the outcome in question is a function of the whole system. On the other hand, the fact that one cannot make the neat conclusions just stated does not call for belittling the experimentation to determine whether retinal location as it relates to one cerebral hemisphere or the other is a differential factor in producing the outcome in question. The task of understanding vision seems to be one of fitting together the various sorts of findings and being guided by knowledge on hand in the further development of expectations and concepts regarding the nature of vision.

The kind of experimentation which tells us most specifically about what the retina does is neurophysiological, and is exemplified by such efforts as the work of Hartline, Granit, and others. In such investigations, it is true, one is not dealing with brightness vision itself, but with the order of functional spatial interrelations which would indicate the *nature* of the discriminating mechanisms to be found in the retina.

Finally what we try to do is, as was indicated above, to put together the various kinds of findings into a unitary picture. In the present case, the effort has been to start with certain basic findings such as provided by the *Limulus* studies and, using essentially the same impingement (stimulus) conditions as used in them, determine what sorts of relations or similarity exist between the two orders of investigation involved.

The technique in the present study involves the presentation of a test target and an agent of brightness change, the inducing target, to one eye and a comparison or match target to the opposite eye. The observer is required to make brightness matches between the test and comparison targets in both the presence and absence of the inducing target; the discrepancy in matches made under these conditions serving as the index of test target brightness change. It was decided to examine the course of this brightness change as the post chiasmal correlates of the inducing and test targets were gradually shifted from the same side to opposite sides. The finding of little or no difference in the effect of the inducing target upon the test target as this manipulation is made, while stimulating a single eye, could be regarded as contributing evidence for the capacity of the periphery to sustain contrast; the finding of large differences could be regarded as contributing evidence against this capacity. Such a shift in central representation could be accomplished by means of an inducing target moving laterally across the visual field, thereby shifting its image from nasal to temporal retina and its central representation in accordance with the pattern of decussation at the chiasma.

B. METHOD AND PROCEDURE

Figure 1 depicts the apparatus employed, with the observer's view of the situation given in Figure 2. The photic source for the match target was a 60 watt incandescent lamp. The optical path included passage through opal glass, a neutral density filter, a stationary prism, a rotary prism, and a two millimeter artificial entrance pupil to the left eye. It was possible for the observer to adjust the position of the artificial pupil to coincide with his own pupil. The photic source for the test target was a 200 watt incandescent lamp. The optical path included passage through opal glass, a neutral density filter, a piece of stationary polaroid J-film, a stationary prism, a rotatable polaroid J-filter, and a two millimeter artificial entrance

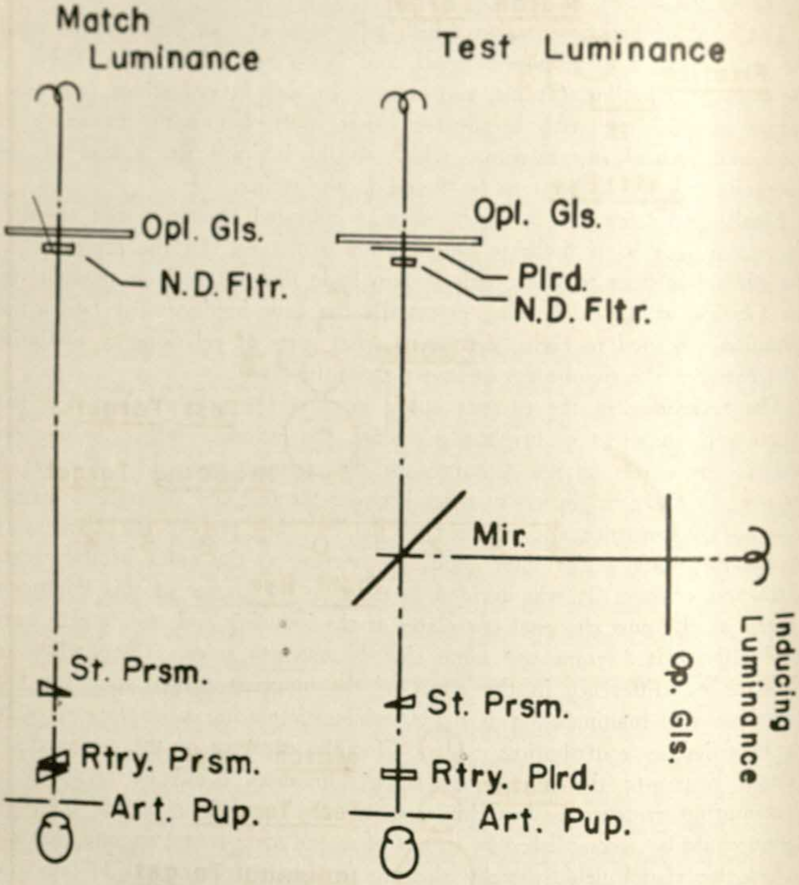


FIGURE 1
Schematic Representation of the Apparatus

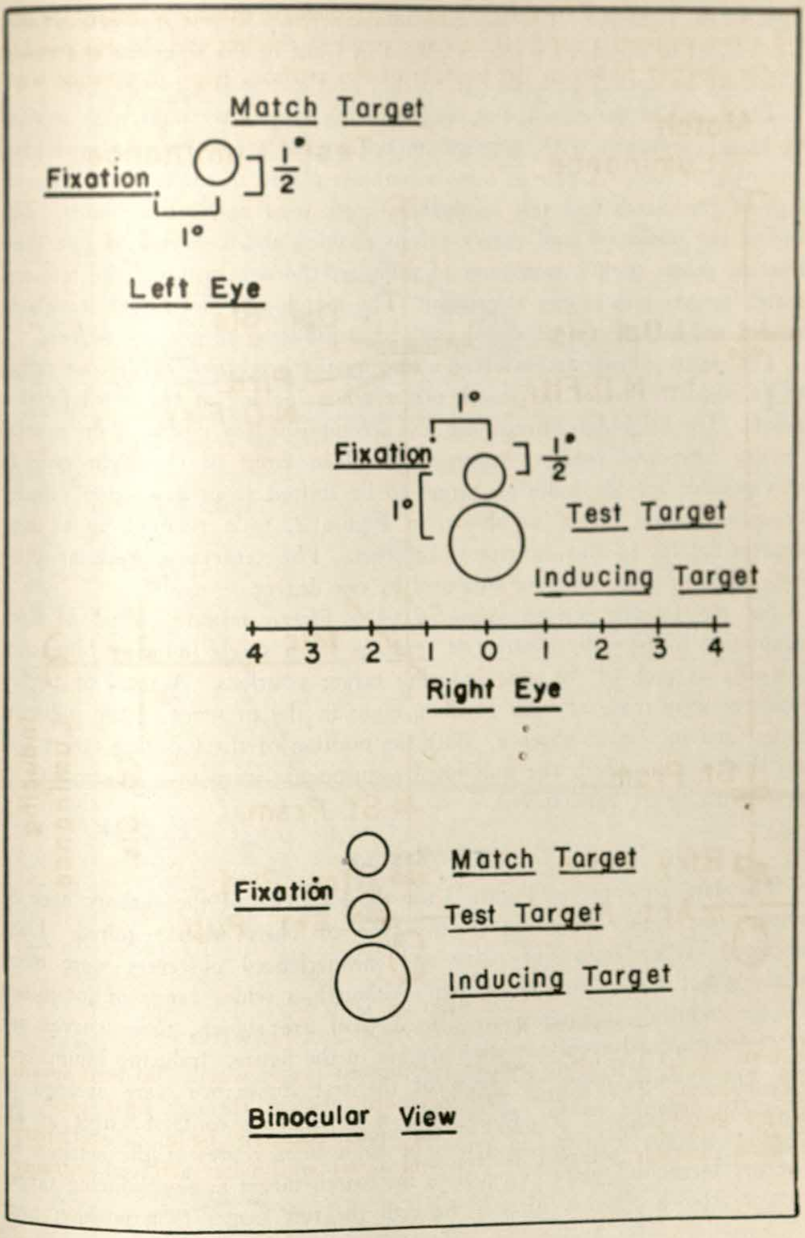


FIGURE 2
Observer's View of the Targets

pupil to the right eye. The inducing luminance, a 750 watt projection lamp, was viewed by means of a mirror placed in front of the right eye and thence following an optical path similar to that of the test luminance.

The forms of the match, test, and inducing targets were given by sections of black cardboard with appropriate openings placed over the opal glass screening of each. A pair of minute pinholes placed in the cardboard shieldings of the match and test luminances were used as fixation points. The use of the stationary and rotary prisms enabled the observers to fuse these fixation points with a minimum of diffiulty, thereby bringing the test and match targets into proper alignment. The match and test targets each subtended one-half degree of visual angle; the inducing target one degree.

The observational task involved adjusting the brightness of the test target to be equal to that of the match target while looking at the fused fixation points. The brightness adjustment was accomplished by means of the neutral density filter and rotating polaroid placed in front of the right eye. It was possible for the inducing target to be shifted from a position directly beneath the test target, as shown in Figure 2, to a position up to four degrees lateral to this in either direction. The centers of each of these nine adjacent positions were separated by one degree.

An experimental session consisted of a fifteen minute period of dark adaptation followed by a series of readings for a single inducing luminance intensity at each of the nine inducing target positions. A total of twelve readings were made at each position, eight in the presence of the inducing target and four in its absence. Both the position of the inducing target and the direction in which the individual adjustments were made, ascending or descending, were randomized.

C. Results

The data presented in Figure 3 are those of two of the authors, each of whom was highly practiced in the type of observation required. Data obtained earlier from five naive and inexperienced observers were more variable but similar to that shown. Although a wider range of intensities of the inducing luminance were used than are shown, these curves are representative and avoid excessive overlap in the figure. Inducing luminances not substantially above the level of the test luminance were avoided in order to produce, as far as possible, a measurable contrast effect at all positions of the inducing target. The solid lines represent the settings of the test target judged to be equal to the match target as the inducing target moved from a position directly beneath the test target to a position four

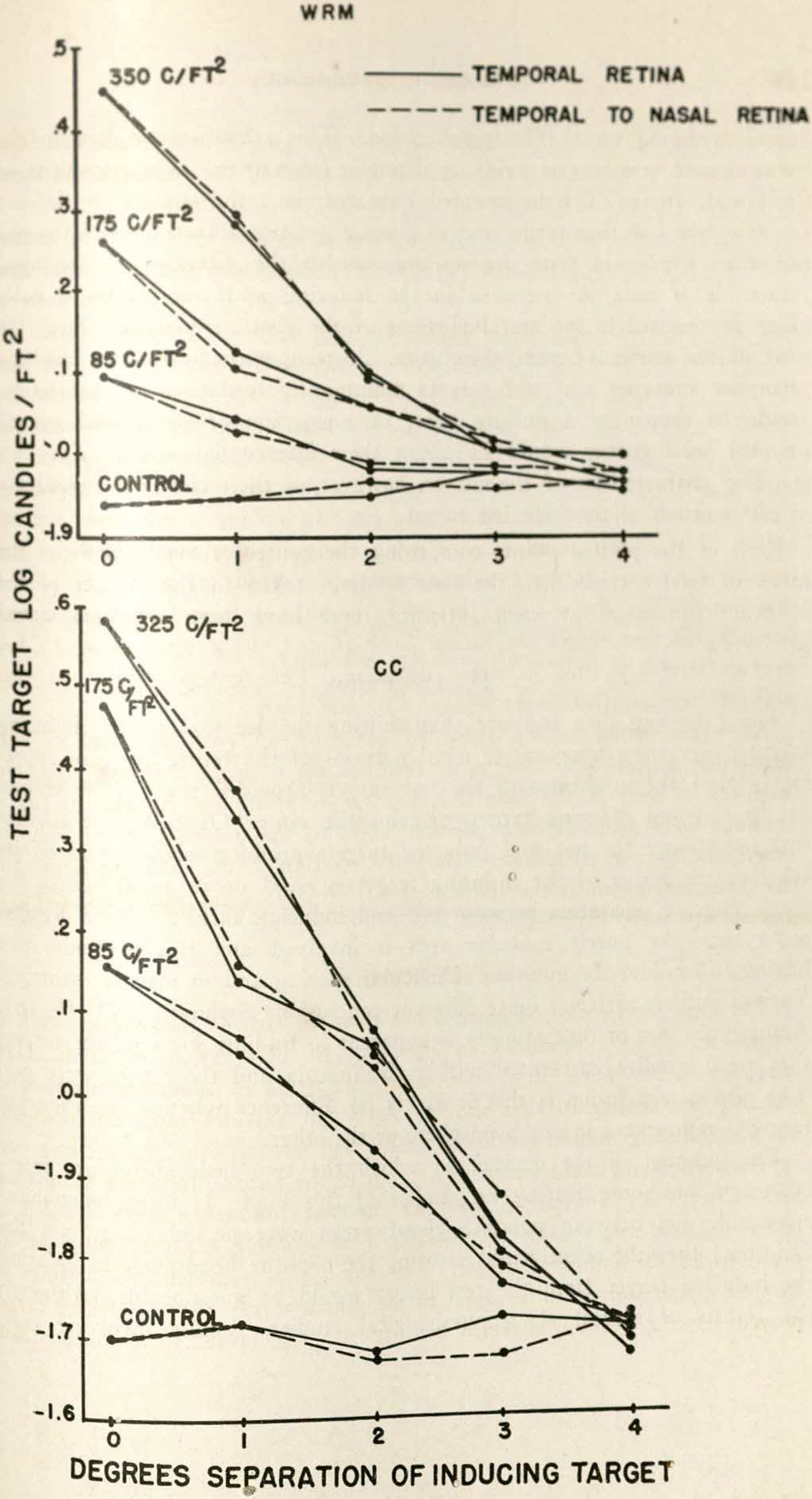

WRM
TEMPORAL RETINA
TEMPORAL TO NASAL RETINA
350 C/FT²
175 C/FT²
85 C/FT²
CONTROL
.5
.4
.3
.2
.1
.0
-1.9
0
1
2
3
4
.6
325 C/FT²
CC
175 C/FT²
85 C/FT²
CONTROL
.5
.4
.3
.2
.1
.0
-1.9
-1.8
-1.7
-1.6
0
1
2
3
4
TEST TARGET LOG CANDLES / FT²
DEGREES SEPARATION OF INDUCING TARGET

FIGURE 6

degrees removed on the temporal retina. This condition resulted in the inducing and test targets being ipsilaterally represented in all positions of the visual system. The interrupted lines represent the settings of the test target as the inducing target moved from a position directly under the test target to a position four degrees removed in the direction of the nasal retina. This condition resulted in the inducing and test targets initially being represented in ipsilateral portions of the visual system and finally in contralateral portions beyond the chiasm. A stronger contrast effect, resulting from the inducing and test targets maintaining ipsilateral representation, would be shown by a pulling apart or separation of the paired curves, i.e., the solid curves would lie above their interrupted mates. The outstanding characteristic of the curves, however, is their close correspondence at all positions of the inducing target.

Each of the plotted points comprising the control curves represents the mean of twelve trials, i.e., the four readings taken in the absence of the inducing luminance at each intensity level have been averaged across intensities.

D. Discussion

The preceding data indicate that shifting the location of the inducing target image from temporal to nasal portions of the retina does not appreciably alter the brightness of the test target, provided that the separation of the test and inducing targets remains the same. Or to put it another way, separating the test and inducing targets produces about equal effects whether the image of the inducing target is nasal or temporal. Since the total range of separation between test and inducing targets is four degrees, more than the purely macular area is involved and this precludes our having to consider the question of macular sparing and/or macular splitting. Various authors arrive at quite different conclusions anyhow (8, 11, 14, 16). Actually no sort of discontinuity is apparent in findings between the targets that were wholly represented within the macula and those that were not. The general conclusion is that it makes no difference whether the inducing target is represented in one hemisphere or the other.

The question of the manner in which the two hemispheres are interconnected has some bearing on studies of this kind. If direct anatomical communication between opposite visual areas were possible (e.g., via the callosum) then the effectiveness of using the opposite hemisphere in isolating the inducing target from the test target would be questionable. Although the results of several electrophysiological studies, usually involving cats, have converged in such a way as to lead to the statement that ". . . almost all individual cortical loci are reciprocally related to points that are symmetrically placed on the opposite hemisphere, as though in mirror image of each other" (15), there is a good deal of evidence to indicate that when dealing with interhemispheric striate area connections in primates this generalization requires qualification (5, 8, 16). Psychophysical evidence can be cited in support of this statement as well. In a control experiment involving the presentation of an inducing target and a match target to opposite eyes in such a way as to be represented in contralateral visual areas, Diamond (7) found that varying the apparent separation between these two targets had an insignificant effect upon the brightness of the match target. There would appear, therefore, to be some agreement between the psychophysical and physiological data on the matter of interhemispheric striate area connections in man and the higher animals.

E. SUMMARY

The effect of an inducing target upon the brightness of a test target was determined by means of a binocular matching technique at each of several locations of the inducing target. At certain locations of the inducing target, the test and inducing targets were represented in ipsilateral portions of the visual system; at other locations the targets assumed contralateral representation beyond the level of the optic chiasma. These shifts in representation as such had no effect upon the brightness of the test target. This demonstration can be taken as evidence for the capacity of the periphery to structure the crucial conditions for brightness contrast, particularly in light of the work of Hartline and colleagues.

REFERENCES

1. Alpern, M. Metacontrast: Historical introduction. *Amer. J. Optom.*, 1952, **29**, 631-646.
2. ———. Metacontrast. *J. Opt. Soc. Amer.*, 1953, **43**, 648-657.
3. Asher, H. Contrast in the eye and in the brain. *Brit. J. Psychol.*, 1950, **49**, 187-194.
4. Bartley, S. H. Vision: A study of its basis. New York: D. Van Nostrand, 1941, Chapter 10.
5. Curtis, H. J. Intercortical connections of corpus callosum as indicated by evoked potentials. *J. Neurophysiol.*, 1940, **3**, 407-413.
6. Diamond, A. L. Foveal simultaneous brightness contrast as a function of inducing- and test-field luminances. *J. Exp. Psychol.*, 1953, **45**, 304-314.
7. ———. Foveal simultaneous brightness contrast as a function of inducing-field area. *J. Exp. Psychol.*, 1955, **50**, 144-152.

8. Duke-Elder, W. S. *Textbook of ophthalmology.* Vol. I. St. Louis: C. V. Mosby, 1942.
9. Fry, G. A., & Alpern, M. The effect of a peripheral glare source upon the apparent brightness of an object. *J. Opt. Soc. Amer.,* 1953, **43**, 189-195.
10. Graham, C. H. Vision. III. Some neural correlations. In C. Murchison (Ed.), *Handbook of General Psychology.* Worcester: Clark University Press, 1934.
11. Halstead, W. C., Walker, A. E., & Bucy, P. C. Sparing and non-sparing of "macular" vision associated with occipital lobotomy in man. *Arch. Ophthal.,* 1940, **24**, 948-962.
12. Hartline, H. K., Wagner, H. G., & Ratliff, F. Inhibition in the eye of *Limulus. J. Gen. Physiol.,* 1956, **39**, 651-673.
13. Kolers, P. A., & Rosner, B. S. On visual masking (metacontrast): Dichoptic observation. *Amer. J. Psychol.,* 1960, **73**, 2-21.
14. Linksz, A. *Physiology of the eye.* Vol. 2. New York: Grune & Stratton, 1952.
15. Livingston, R. B. Central control of receptors and sensory transmission systems. In C. Field, H. W. Magoun, & V. E. Hall (Eds.), *Handbook of Physiology.* Vol. I, Section I. Washington, D.C.: American Physiological Society, 1959.
16. Polyak, S. P. The vertebrate visual system. H. Kluver (Ed.). Chicago: University of Chicago Press, 1957.
17. Schouten, J. F., & Ornstein, L. S. Measurements of direct and indirect adaptation by means of a binocular method. *J. Opt. Soc. Amer.,* 1939, **29**, 168-182.
18. Sherrington, C. S. On reciprocal action in the retina as studied by means of some rotating discs. *J. Physiol.,* 1897, **21**, 33-54.
19. Toch, H. The perceptual elaboration of stroboscopic presentations. *Amer. J. Psychol.,* 1956, **69**, 345-358.
20. Werner, H. Studies in contour: I. Qualitative analysis. *Amer. J. Psychol.,* 1935, **47**, 40-64.
21. ———. Strobostereoscopic phenomena. *Amer. J. Psychol.,* 1940, **53**, 418-422.

Department of Psychology
Michigan State University
East Lansing, Michigan

$20.00 per annum
$15.00 per volume
$7.50 per half volume

QUARTERLY
Two volumes per year
Immediate publication

October, 1961
Volume 52, Second Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

OCTOBER, 1961

(Manuscripts are printed in the order of final acceptance)

The analysis of mutual concurrences 253
By J. A. Gengerelli

The role of hypnosis in self understanding 265
By Seymour Halpern

A psychoanalytic approach to Kabuki: a study in personality and culture . . 269
By John D. Mitchell and E. K. Schwartz

Psychosomatic V 281
By Herbert C. Archibald, Dorothy Bell, Christine Miller, and Clare W. Thompson

Convergence as a cue to absolute distance 287
By Walter C. Gogel

Convergence as a cue to the perceived distance of objects in a binocular configuration 303
By Walter C. Gogel

(over)

Copyright, 1961, by The Journal Press
Provincetown, Massachusetts, U. S. A.
Entered as second-class matter January 26, 1937, at the post-office at
Provincetown, Mass., under the Act of March 3, 1879
Second-class postage paid at Provincetown, Mass.

Magnetic fields accompanying transmission of nerve impulses in the frog's sciatic 317
By J. A. Gengerelli, Norman J. Holter, and Wilford R. Glasscock

The effect of analgesic dosages of acetylsalicylic acid upon psychological reaction time 327
By Margaret Kathleen Cappone

Assimilation and contrast effects of anchoring stimuli on judgments: a partial replication of the Sherif, Taub, and Hovland study 333
By Lucy Bravo and M. S. Mayzner

Deviations from matching behavior in probability learning 335
By Joan E. Jones

Open-mindedness and hypotheses of interdependence between libertarian and equalitarian processes 347
By Salvatore V. Zagona and William J. MacKinnon

Head balance and sitting posture II: the role of the sternomastoid muscle . 363
By Frank Pierce Jones, John A. Hanson, and Florence E. Gray

Concept-forming behavior in schizophrenic and non-schizophrenic subjects . 369
By Paul McReynolds and Beverly Collins

Reinforcement and personality factors in verbal conditioning 379
By Arnold Binder and Phyllis Salop

Response strength as a function of rate of reward 403
By Ronald L. Jorgensen and Roger T. Davis

Studies on transfer of brightness discrimination from bar-pressing to a Y maze 411
By M. R. D'Amato and H. Jagoda

Volume-unit graphs: their production and applicability in studies of infants' cries 423
By Vincent R. Fisichelli, Samuel Karelitz, John Eichbauer and Laura S. Rosenfeld

Conditions for creative problem solving 429
By L. Richard Hoffman

Lysergic acid diethylamide (LSD-25) XXXIII. Effect on surfacing behavior of the carp and possible application to game conservation and survival kits 445
By H. A. Abramson, H. H. Gettner, M. P. Hewitt, and G. Dean

The roll of PCF in temporal manipulations of color 457
By Thomas M. Nelson and S. Howard Bartley

The effect of areal characteristics of targets upon shape-slant invariance . . 479
By Thomas M. Nelson, S. Howard Bartley and Charles M. Bourassa

Fantasy testing for ESP in a fourth and fifth grade class 491
By M. L. Anderson and R. A. McConnell

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 253-263.

THE ANALYSIS OF MUTUAL CONCURRENCES*[1]

Department of Psychology, University of California

J. A. GENGERELLI

A. GENERAL

If there are two items *a* and *b*, each one of which may occur or not occur, we can determine whether any covariation or concurrence exists between them by forming a 2×2 table and calculating X^2 or some related coefficient. If there are three items, *a, b,* and *c,* and we ask whether covariation exists between the three, we can, as in the bivariate case, calculate the appropriate coefficients for *a b,* and *a c* and *b c;* if each of the three values meets a prespecified level of statistical significance, mutual covariation may be asserted to exist between the three elements. Should one of the coefficients be insufficient, mutual concurrence between *a, b,* and *c* would be denied.

Let us next imagine a set of *n* elements: *a, b,* .... *n,* each one of which may or may not occur. We ask: are there one or more sub-sets in the set *n* whose members mutually covary, and if so, what are the members of these sub-sets? Many of the $_nC_2$ pairs will reach the requisite level of significant covariation, but many certainly will not; since, however, we seek a sub-set of elements, all of whose members covary mutually, our problem becomes vexingly difficult. To determine which of the elements in a set constitute a sub-set whose members are mutually concurrent is the essence of the problem of pattern analysis in psychology. Whether we concern ourselves with clinical syndromes, personality traits, or social attitudes, the task remains the same: does a structure exist? If so, what are its elements?

The method to be described does not attempt to penetrate below the level of empirical phenomenology. Thus, factor analysis, and to a lesser degree, Tryon's method of Cluster Analysis (5), are motivated to find the factors which presumably are responsible for the empirical covariations. Latent structure analysis (1, 2, 3), similarly, seek to disentangle the sub-populations comprising a given population. The present procedure is more restricted in

---

* Received in the Editorial Office on May 10, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Supported in part by a grant from the Holter Research Foundation, Helena, Montana, a nonprofit organization.

that it merely purports to achieve descriptive classification of the data given rather than to yield theoretical structures which may be conceived as generating the data. On the other hand, the method has the virtues of its limitations since fewer assumptions are required in evaluating the results.

The problems encountered by the method are two: (1) the isolation of sub-sets of elements which covary mutually, and (2) determination of the statistical significance of the sub-sets found.

In the interests of simplicity a single paradigm will be used in describing the method. However, the procedure is quite flexible and can be modified to apply to a variety of data; the interested investigator will have no difficulty in adapting the technique to his particular needs. Nor must it be supposed that the method is limited to elements of a dichotomous nature; the "elements" may be in the nature of variables which assume continuous values. In this latter case, the data utilized by the method will be in the nature of Pearsonian or similar coefficients.

B. The Isolation of Sub-Sets of Concurrences

Consider a check-list made up of *m* items describing psychological characteristics, and let each one of these be of an all-or-none nature. Let there be a population of M patients admitted to a neuropsychiatric hospital, and each patient checked on each of the items in the list. A check will be made opposite an item only if the patient manifests the characteristic in question; if the patient is free of the characteristic, no mark is made. Our problem is to determine whether in the population of subjects examined there are, among the *m* characteristics, sub-sets of mutual concurrences; or, in different language, whether certain members of the set *m* form one or more syndromes or constellations.

Construct a square matrix with vertical and horizontal arrays numbered from *1* to *m*. Taking, now, each patient in turn, a mark is made in each cell formed by the intersection of the vertical and horizontal arrays denoting those symptoms found present in the patient. Thus, if a given patient has the characteristics 3, 6, 14, 32, a mark would be made in the cells (3, 6), (3, 14), (3, 32), (6, 14), (6, 32), (14, 32). Since the matrix used is square (rather than a half matrix), there will be two cells for each combination of numbers, and a mark should be made in each. This would, in general, entail $x(x-1)$ marked cells for each patient, where x represents the number of characteristics found in the patient.

When the data from each of the M patients have been so recorded, the cells in the matrix will contain varying numbers of entries. The immediate problem is to find from these entries whether sets of mutual concurrences or associations exist, and we will do this by determining whether the number of entries in each of the $\frac{(m)\ (m-1)}{2}$ cells (in the half-matrix) is significantly different from that expected by chance.

Sum the number of marks in each vertical array; this gives n_j, where j represents the vertical array. If we sum the n_j's and divide by 2 (the data in the matrix have been recorded twice), we will have the number of associations which have occurred among the "symptoms" in the populations studied. Let us designate this quantity N. That is, $N = \sum_j n_j/2$. N could have been computed, of course, on the basis of the horizontal array sums.

The expected frequency (under the null hypothesis) of the marks in the cell formed, for example, by the intersection of the arrays 3 and 4 would be given by $\frac{(n_3)\ (n_4)}{N}$; and similarly for each cell in the matrix. Once the expected frequency for each cell is computed, we proceed to test the deviation of the number of observed entries in a given cell from that expected. This is achieved by breaking down the matrix into a 2 × 2 table and applying X^2 for each cell. For the (3, 4) cell we would have the situation shown in Table 1.

TABLE 1

2 × 2 TABLE REQUIRED IN CALCULATING CHI SQUARE FOR CELL (3, 4)

	3	not 3	
4	(A) f_{34}	(B) $f_4 - f_{34}$	f_4
not 4	(C) $f_3 - f_{34}$	(D) $(N - f_3) - (f_4 - f_{34})$	$N - f_4$
	f_3	$N - f_3$	

In A is recorded the number of entries in cell (3, 4); in B, the number of entries in array 4 minus those in cell (3, 4); in C, the number of entries in array 3 minus those in cell (3, 4); in D, the difference between $(N—f_3)$ and

$(f_4—f_{34})$. Since we have already calculated the "expected" value for cell (3, 4), we can without further computation supply the respective expected values for the other three, and proceed to calculate X^2 for the table. The significance level should be made rather stringent, in order to keep the number of chance associations in the matrix small. The .01 level is recommended. The usefulness of this maneuver will become evident later.

When the above has been done for each cell in the matrix[2], some cells will have significant X^2, others, probably the great majority, will not. The "significant" cells may then be marked, to make them easily distinguishable from the others. Our task now is to determine whether the marked cells contain among them "constellations" or "syndromes." For example, if there were a syndrome or "concurrence sub-set" constituted of the symptoms 2, 6, 13, 28, 30, we would find the following cells to be marked: (2, 6), (2, 13), (2, 28), (2, 30), (6, 13), (6, 28), (6, 30), (13, 28), (13, 30), (28, 30).

Sum the number of significant cells in each vertical array of the square matrix [see Table 2]. To each sum add the digit "1." The augmented sum for a given vertical array represents the maximal size of the concurrence constellation of which that symptom could be a member. Choose the two vertical arrays having the largest sums. Let us suppose these to be columns 1 and 10. Check the (1, 10) cell in the matrix, to determine whether it

TABLE 2

SYNTHETIC TABLE REPRESENTING THE SET OF CONCURRENCES AMONG 14 ITEMS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
1		O	X	O	X	X	O	X	O	X	O	O	O	O	6
2	O		O	O	O	O	X	O	O	O	O	O	O	O	2
3	X	O		O	X	X	O	O	O	X	O	O	X	O	6
4	O	O	O		O	O	X	X	O	O	O	O	O	O	3
5	X	O	X	O		X	O	O	O	X	O	O	O	X	6
6	X	O	X	O	X		O	O	O	X	O	X	O	O	6
7	O	X	O	X	O	O		X	X	O	O	O	X	O	6
8	X	O	O	X	O	O	X		X	O	O	O	O	O	5
9	O	O	O	O	O	O	X	X		O	X	O	O	O	4
10	X	O	X	O	X	X	O	O	O		O	X	O	X	7
11	O	O	O	O	O	O	O	O	X	O		O	O	O	2
12	O	O	O	O	O	X	O	O	O	X	O		O	O	3
13	O	O	X	O	O	O	X	O	O	O	O	O		O	3
14	O	O	O	O	X	O	O	O	O	X	O	O	O		3
	6	2	6	3	6	6	6	5	4	7	2	3	3	3	

The cells marked *X* are significant cells: those marked *O* are not significant.

---

[2] Most of the work can be done by inspection once the expected frequencies are computed.

is a significant cell. Let us suppose (1, 10) is significant. We next choose the column having the next largest sum, say, 7. We now compare 7 with columns 1 and 10. If item 7 is mutually concurrent with items 1 and 10, cells (1, 7) and (10, 7) should be significant cells. Let us suppose one of them is not. (Table 2). We therefore discard 7. The next largest sum belongs to item 5. We now check cells (1, 5) and (10, 5). Both are significant cells. We therefore conclude that items 1, 10 and 5 are mutually concurrent. Proceeding in this manner, we find (Table 2) that items 1, 3, 5, 6, 10 form a set of mutually concurrent symptoms and therefore constitute a constellation.

It will be noted (Table 2) that items 2, 4, 7, 8, 9, 11, 12, 13 and 14 remain. Since of these items 7 and 8 have the two largest sums, we examine cell (7, 8) for significance; it is significant. Item 9 has the next largest sum. Examining cells (9, 7) and (9, 8) we find them both significant. We have, therefore, a second constellation constituted of items 7-8-9. No other item can be a member of this constellation of mutual occurrences since no other item shows a sum greater than 3, while a sum of 4 is required.

We have, therefore, found two constellations in our set of 14 items, one constituted of 5 elements, (1-3-5-6-10), the other of 3, (7-8-9).

The above inspection procedure, barring proofreader error, will always reveal such constellations as exist in the matrix; and accomplishes this quite simply.

C. The Statistical Significance of Concurrences

There now remains the problem of the statistical significance of a sub-set of concurrences. Clearly if in the $m \times m$ matrix m is a large number, and, in addition, a considerable number of its cells are significant cells, chance alone would provide a constellation involving, say, three elements. What is required, therefore, is a probability statement for each sub-set of concurrences found in the matrix.

The frame of thought is that there are but two givens: (1) the size of the matrix used and (2) the number of significant cells. It is as if a blindfolded marksman were to throw simultaneously a handful of darts, d in number, at a half-matrix of K cells, with the restriction that each dart would hit one of the cells, that no cell would be hit by more than one dart, and that one cell is as likely to be hit as another. We are to evaluate the outcome of the event on the assumption that the loci of hits are mutually independent.

Consider a half-matrix with $(m)(m-1)/2 = K$ cells. Let there be *d* significant cells in the matrix (i.e., cells which fulfill the requirements of the X^2 test). For the moment, let us subject the number *d* to the restriction that it be equal to ${}_tC_2$, that is, the number of combinations of *t* things taken two at a time; also $d << K$. These *d* significant cells could be distributed in $D = {}_KC_d$ ways or patterns on the half-matrix. It is essential that we divide patterns into two mutually exclusive classes: *closed* patterns and *open* patterns. A *closed* pattern is one, all of whose elements are mutually concurrent. For example, if we are considering a pattern which involves the elements *m, n, o, p* and the cells (m, n), (m, o), (m, p), (n, o), (n, p), and (o, p) are significant cells, the pattern is said to be a *closed* pattern. Hence, if elements involved in the *d* significant cells behave in a mutually concurrent manner, the pattern will be called a *closed* pattern. If, however, not *all* elements involved in the *d* significant cells behave in this mutually concurrent manner, the set of cells will be termed an *open* pattern.

The total number of possible patterns of *d* cells each, both open and closed, in the half-matrix is D; but, as is intuitively evident, the number of possible closed patterns will be much smaller. It will be necessary to determine for an $m \times m$ half-matrix how many closed patterns constituted of *d* cells are possible.

The closed pattern of *d* cells is constituted of, let us say, *t* elements. We can conceive of the closed pattern as a half-square with the sides $t \times t$. The first column is of (t-1) cells, the second column of (t-2) cells, the third, of (t-3) cells; etc. Turn to Table 2. Imagine that the (t-1) cells in the first column of the pattern coincide with the first (t-1) cells in the first column of the half-matrix; the other cells in the pattern will be superimposed on the corresponding cells in the matrix. This is one possible position. With the restriction that the first column of the pattern shall coincide with the first column of the matrix, ${}_{m-1}C_{t-1}$ arrangements are possible altogether. Next, imagine the pattern so placed that its first column is superimposed upon the second column of the matrix. Within this restriction, ${}_{m-2}C_{t-1}$ arrangements are possible. Proceeding in this manner, we will have, altogether $({}_{m-1}C_{t-1} + {}_{m-2}C_{t-1} + {}_{m-3}C_{t-1} + \cdots + {}_{t-1}C_{t-1})$ total possible arrangements or places which the closed pattern of *d* elements can occupy on the $m \times m$ half-matrix. In different language, this sum represents the number of closed patterns of *d* cells which can appear on the matrix. Call this quantity E. Fortunately, we can obviate the above very tedious summation, for it can be shown that the sum is given by ${}_mC_t$. That is, $E = {}_mC_t$.

Since the total number of possible patterns on the half-matrix, whether open or closed, is D, then the probability that a pattern of the specified size is a *closed* pattern will be given by $p_E = E/D$, and the probability that a pattern of that size will be an *open* pattern will be $p_F = 1\text{-}p_E$.

If, therefore, in an experiment of the type described we obtained only *d* significant cells, and these formed a *closed* pattern when *m* items had been tested, the probability of obtaining such a sub-set of mutual concurrences by chance would be given by the ratio

$$p_E = {}_mC_t/{}_kC_d = E/D \qquad [1]$$

If this ratio is sufficiently small, we would say that the constellation of symptoms is "real."

Unfortunately, the number of significant cells to be found in the half-matrix will usually exceed the number of cells forming a closed pattern or constellation. Thus, in Table 2 there are 24 significant cells, but the largest closed pattern is constituted of only 10 cells. How does this affect the probability that the closed pattern in question was an accident?

Let the number of significant cells in the half-matrix be h, when $h > d$. The question now becomes: What is the probability of securing at least one *closed* pattern, constituted of d cells, when h significant cells are to be found in an $m \times m$ half-matrix? A number of patterns of size d can be made from a set of h elements. The number of such patterns is given by ${}_hC_d$. Some of these patterns will be closed patterns, some open. Now, the probability for a closed pattern of d cells in the $m \times m$ half-matrix is p_E and the probability of an open pattern is p_F. Therefore, the probability of forming 0, 1, 2, 3, ... ${}_hC_d$ *closed* patterns from the set of h significant cells is given, respectively, by the terms in the binomial expansion

$$(p_E + p_F)^{{}_hC_d}$$

The evaluation of these terms is not an insignificant computational chore. However, we need to evaluate only the first term, since we seek the probability of obtaining at least one closed pattern. Let ${}_hC_d = v$. The first term is then given by ${}_vC_o\, p_E^o p_F^v$, which yields the probability that none of the patterns will be *closed* and all *open*. But ${}_vC_o p_E^o p_F^v = p_F^v$. Therefore the probability that there will be at least one closed pattern in the set of patterns is given by $1 - p_F^v$. If we designate this quantity by the symbol P_E, the equation

$$P_E = 1 - p_F^v \qquad [2]$$

is the basic equation sought.

From this equation it is readily seen that the significance of the closed pattern obtained is determined by (*a*) the probability of obtaining an open pattern of a given number of cells in a half-matrix of a specified size and (*b*) the number of patterns, open or closed, of the specified size which can be formed from the number of significant cells occurring in the matrix. Clearly, with all else fixed, as the number of significant cells increases, v increases, hence the size of P_E increases, i.e., the level of significance of the obtained closed pattern diminishes.

D. ILLUSTRATIONS AND DISCUSSION

By way of illustration, computations will be made with the data shown in Table 2. Since the table is 14×14, the number of available cells is $K = (14)\ (13)/2 = 91$. We will determine the significance of the largest closed constellation obtained. Since this pattern involves 5 elements (t), the number of cells involved is $d = (5)(4)/2 = 10$. Ten significant cells can be distributed over the half-matrix in $D = {}_{91}C_{10} = 6.427 \times 10^{12}$ different ways. That is, there are possible 6.427×10^{12} distinguishable patterns, either open or closed.[3] The number of possible closed patterns is given by ${}_{14}C_5 = E = 2002$. The probability that 10 significant cells in the matrix in question will form a closed pattern is given by

$$p_E = E/D = 2002/6.427 \times 10^{12} = .0000000003$$

The probability that they will form an open pattern is given by

$$p_F = 1\text{-}p_E = .9999999997$$

Since there are 24 significant cells in Table 2, they are capable of yielding ${}_{24}C_{10}$ patterns, open or closed, of 10 elements; that is, 1,961,132 patterns. The probability that 24 significant cells in a half-matrix of 91 cells are capable of yielding at least one closed pattern is given by subtracting the value of the first term in the binomial $(p_E + p_F)^{{}_{24}C_{10}}$ from 1. This is given by the formula

$$P_E = 1 — p^{{}_{24}C_{10}}$$

Since the numbers now become too unwieldy, we replace them by approximations, and have

$$\begin{aligned} P_E &= 1—(.999999999)^{2 \times 10^6} = 1—.9994 \\ &= .0006 \end{aligned}$$

---

[3] As the need arises, we will substitute approximate values for the actual ones.

A far simpler procedure for the last phase of the calculation is to utilize a table giving the cumulative sum of the Poisson distribution (4, Table 7, p. 122). All that is required for using the table is the mean for the binomial. This is given by $m = (P_E)({}_hC_d)$. In the problem being considered $m = .0006$.[4] The tabled value for the first term of a Poisson when $m = .0005$ is given as .9995. Therefore, $P_E = 1—.9995 = .0005$, which is in sufficiently close agreement with the result of the more direct calculation.

Our closed pattern thus is significant, and we conclude that in the set of 14 elements there is a subset of 5 elements which examplify mutual concurrence.

The other closed pattern in Table 2 is of 3 elements. Determination of the "reality" of this pattern introduces a new consideration not met previously. We have already shown the reliability of the larger pattern of 5 elements. Had this larger closed pattern not met the criteria of statistical significance, we would know *ipso facto* that the smaller would certainly not do so. Since, however, the 5-element pattern is real, our analysis of the requisite probabilities are altered. By virtue of the fact that certain cells are no longer available, we cannot utilize the values of p_E and p_F already calculated. We need to make a new analysis.

Turning to Table 2, we strike out columns (and rows) 1, 3, 5, 6, 10. This now leaves us with a reduced 9 × 9 matrix, and with but 8 significant cells.

The number of available cells in the half-matrix is now $K = \frac{(9)(8)}{2} = 36$. Since the closed pattern in question is of 3 elements, the number of cells involved is $d = \frac{(3)(2)}{2} = 3$. Three significant cells can be distributed over the half matrix in $D = {}_{36}C_3 = 7152$ different ways. The number of possible closed patterns is given by ${}_9C_3 = 84$. The probability that 3 significant cells will form a closed pattern is given by

$$p_E = \frac{E}{D} = \frac{84}{7152} = .0017$$

The probability that they will form an open pattern is given by

$$p_F = 1 - p_E = 1 - .0017 = .9983$$

---

[4] The symbol *m* here denotes the mean of the Poisson (i.e., binomial), and is not to be confused with our previous use of the same symbol.

Since there are 8 significant cells in the reduced matrix, they are capable of yielding ${}_8C_3$ patterns, open or closed, of 3 elements; that is, 56 patterns. The probability that 8 significant cells in the existing half-matrix are capable of yielding at least one closed pattern generated by 3 elements is given by

$$P_E = 1 - p_F^{{}_8C_3} = 1 - (.9983)^{56} = 1 - .909 = .091$$

If we utilize the Poisson function, we have

$$m = (p_E)\ ({}_8C_3) = (.0017)\ (56) = .0952$$

The tabled value for the first term of a Poisson when $m = .10$ is given as .9048. Hence

$$P_E = 1 - .9048 = .0952$$

Our second 3-element pattern is, therefore, not significant at the .05 level.

The computational requirements for determining the "reality" of constellations which turn up, while heavy, are not prohibitive. Considerable work is avoided if the following rule of thumb is observed. After p_F and ${}_hC_d$ (the number of patterns, open or closed, of *d* cells which may be formed from *h* significant cells) have been determined, express p_F as a fraction, and note the order of magnitude of its denominator. In our example of the 5-unit constellation, this was 10^{10}. Then note the order of magnitude of ${}_hC_d$; in our example, again, this was 10^6. Unless the order of magnitude of the denominator of p_F is larger than that of ${}_hC_d$, it will not be worth while to continue the computation, since P_E will be too large. The denominator of the fraction should be at least 2 orders of magnitude larger than that of ${}_hC_d$. It is preferable to avoid this phase of the computation altogether by using a Poisson table.

The method which has been described seems to hold some promise for attacking certain problems which arise in psychology and related disciplines. In the illustration given, it would seem that the question of nosological classification may be profitably approached from this point of view. What pathological *entities* may be said to exist, if any, and what are the elements comprising each? Constellation of attitudes in a population of persons, and of cultural characteristics in a population of communities may be similarly isolated and analyzed.

E. SUMMARY

A method is presented for determining the existence of mutual concurrences in a subset of events forming part of a larger set.

Formulae are derived to determine the statistical significance of patterns which may be found.

An illustration is given, in terms of a synthetic clinical problem, of the computational procedures required.

REFERENCES

1. ANDERSON, T. W. On estimation of parameters in latent structure analysis. *Psychometrika,* 1954, **19**, 1-10.
2. GREEN, B. A general solution for the latent class model of latent structure analysis. *Psychometrika,* 1951, **16**, 151-166.
3. LAZARSFELD, P. F. The use of mathematical models in the measurement of attitudes. *Research Memorandum RM-455,* Rand Corp., March 22, 1951.
4. PEARSON, E. S., & HARTLY, H. O. Biometrika tables for statisticians. Vol. 1. 1956. London: Cambridge University Press.
5. TRYON, R. C. Cumulative community cluster analysis. *Ed. & Psychol. Meas.,* 1958, **18**, 3-35.

Department of Psychology
University of California
Los Angeles 24, California

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 265-267.

THE ROLE OF HYPNOSIS IN SELF UNDERSTANDING*

Central Islip State Hospital, New York

SEYMOUR HALPERN

Within the past two decades, hypnosis has once again emerged in the field of legitimate scientific enquiry (3, 5). Its endorsement by the Council on Mental Health of the American Medical Association is further evidence of its current favorable status (1). Nevertheless, there still exists among many practitioners and social scientists an attitude of reserve and scepticism concerning the application and limitations of hypnosis. These attitudes, however, have not significantly dampened the increasing interest which might well be construed as an expression of the contemporary Zeitgeist.

Curiously enough the resurgence of interest in hypnosis has been paralleled by an increased number of publications concerning the problems of consciousness. Confronted with an unprecedented threat to existence, we apparently are looking inwards with greater scrutiny for solutions to the vast social problems which have been created by our irrational strivings.

Hypnosis is an invaluable introspective capacity for the exploration of one of nature's most profound mysteries, consciousness. These problems are most acutely felt by those laboring in the field of mental health. Self-knowledge is no longer a matter of wisdom, it has become a desideratum of great urgency. Despite their many shortcomings, often stemming from rigid doctrinaire attitudes, psychoanalysts have been in the fore among those seeking greater self-awareness. It would be unthinkable for an analyst to undertake the treatment of a patient without prior personal analytical experience. Incidentally, this has been a major factor in the promotion of professional competency among psychotherapists. Consequently, medical specialists employing hypnosis have been enjoined to acquire the rudiments of psychodynamics as a safeguard against unintentional "malpractice." This plea is a modern echo of the ancient precept, "physician heal thyself." When human existence is considered a process, the evolution of self-knowledge may be seen as an interminable quest. Let us consider the role of hypnosis in this quest. Psychoanalysis has clearly demonstrated the value of the introspective

---

* Received in the Editorial Office on May 15, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

method when its purpose is related to self-growth and the social process. In a symposium on problems of consciousness, David Rapaport, one of the leading psychoanalytic theorists, commented on his own method of self-exploration through the analysis of hypnogogic imagery (4). Certainly a practice of a master ought to be considered by a student. In a previous communication on the relationship between hypnotically induced percepts and psychodynamics, the author touched upon the same problem (2). The capacity for self-understanding is cogently demonstrated by the hypnoanalytic interpretation of dreams; many of the author's patients have shown the ability to arrive at adequate interpretations of their own dreams. It is a practice of the author to employ auto-hypnosis for the exploration of psychodynamics. This procedure might be compared to the meditations of the religious. It has been reported that the hypnotic interaction is enhanced when the therapist concurrently induces an auto-hypnotic state.[1]

Whereas man grows old effortlessly, psychological maturity or wisdom is only acquired as the result of a never-ending struggle against the forces of inertia. This universal resistance to change is nowhere as dramatically demonstrated as during the psychotherapeutic process. The hypnotherapist by the use of hypno-introspection is in a singular position to further his own self-growth and acquire insights into the nature of consciousness. In addition to subserving the elucidation of the transference-countertransference continuum, this procedure of self-evaluation might well contribute to the growth of the conviction within the therapist that his relationship to the patient is that of collaborator and not manipulator, an assertion to which much lip-service is often made. Hypno-introspection should also enhance the hypnotist's appreciation of the experiences of his patients. Furthermore, the knowledge of the capacity for self-development serves to encourage wholesome inter-dependence, thus mitigating the forces of infantilization.

REFERENCES

1. Council on Mental Health: Medical Use of Hypnosis, J.A.M.A., 168: 186-189, 1958.
2. HALPERN, S. On the Similarity between Hypnotic and Mescaline Hallucinations: A paper delivered at the 12th Annual Meeting of the Society for Clinical and Experimental Hypnosis on October 6, 7, 1960, in Washington, D.C.
3. KLINE, M. V. (*Ed.*) Hypnodynamic Psychology. The Julian Press, Inc., New York, 1955.

---

[1] Communication by Dr. Margaretta Bowers during a workshop at the 12th Annual Meeting of the Society for Clinical and Experimental Hypnosis in Washington, D.C. on Oct. 5, 1960.

4. Rapaport, D. Consciousness, a psychopathological and psychodynamic view. In *Problems of Consciousness, Transactions of the Second Conference,* March 19-20, 1951. (Ed. by H. A. Abramson.)

5. Schneck, J. M. (*Ed.*) Hypnosis in Modern Medicine. Charles C Thomas, Springfield, Ill., 1953.

Research Division
Central Islip State Hospital
Central Islip, New York

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 269-280.

A PSYCHOANALYTIC APPROACH TO KABUKI: A STUDY IN PERSONALITY AND CULTURE*

Institute for Advanced Studies in the Theatre Arts, New York City
and Postgraduate Center for Psychotherapy, New York City

JOHN D. MITCHELL AND E. K. SCHWARTZ

Wide interest in things Asian, especially Japanese, has overtaken the United States in the last two decades. Since World War II anything Japanese is attractive to Americans, who tend to idealize the Japanese and their culture, perhaps as a reaction formation to the great hostility for them and from them during the most deadly days of combat. Americans have even become deeply involved in products which the Japanese have long since largely discarded, Zen Buddhism, for example.

In the Fall of 1960 in New York, Onoe Baiko VII of Japan directed non-Japanese professional actors in the first production in English of a classical Kabuki play.[1] Baiko was an international guest and a visiting director for an All American troupe which performed Kabuki for invited professional audiences at the Institute for Advanced Studies in the Theatre Arts in New York City.[2]

Kabuki, like all great traditional theatres is grounded in a view of life and a way of living. Kabuki is universal and yet profoundly Japanese. The primary purpose for American actors performing a Kabuki play under the direction of a great Japanese master of the *Kata* (style) of Kabuki was

---

* Received in the Editorial Office on May 22, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Onoe Baiko VII, famed *onnagata* and actor-dancer of classic roles in the Kabuki theatre, is the eldest son of the late Onoe Kikugoro VI. At the age of five, as Onoe Ushinosuke IV, Baiko made his stage debut in 1920. In addition to his Kabuki training, he received a university education. In 1935 he was given the name of Kikunosuke II and began to build a reputation for himself as an *onnagata* actor (one who plays female roles). He succeeded to the distinguished title Baiko in 1947 and became a leader in the Kikugoro Troupe in 1949. He is one of the most versatile artists in the Kabuki theatre and had never before left his native land. He has distinguished himself not only in *onnagata* roles but also in young hero roles and as a solo dancer.

[2] The Institute for Advanced Studies in the Theatre Arts (IASTA) is an educational organization dedicated to the enrichment of the American theatre through greater awareness of the achievements, past and present, of world theatre. The Institute brings to New York outstanding theatre practitioners from all over the world to work with interested and qualified American actors, directors, playwrights and scenic artists, who are enrolled as Fellows of the Institute. This enables the American theatre artist to study at first hand the styles and techniques of foreign traditions in his own language and without leaving home.

to provide them with new perspectives and possible applications in their own Western theatre.

Mastery of the style or patterns (*kata*) of Kabuki acting begins in Japan at the age of five and requires an apprenticeship which may last forty years. The challenge to both the director and his American actors was to capture something of the spirit and aesthetics of the four-hundred-year-old tradition. What took the audiences by storm was not alone a remarkable demonstration of American actors' mastery of a fantastic technique of acting, but also the assault on the unconscious by both the Kabuki style and the content of the particular play, *Narukami. Narukami* is a typical Kabuki and belongs to *Juhachiban,* a group of eighteen plays which were the "hits" of the great acting family of Ichikawa Danjuro. It was first performed in February, 1684.

The theme of the play is taken from the Noh drama *Ikkaku Sennin,* the name of a priest of India who was ruined by a woman and sake. The play remains a perennial favorite, for its theme is the universal problem of human weakness. In the Kabuki version the priest Narukami has been piqued and angered by the Imperial Court. For revenge he has used his powers to imprison the Rain God. When the play begins, drought is upon the land.

In a mountain retreat, Narukami and his acolytes are practicing ritual austerities. Sent by the Imperial Court the young princess Taema arrives, claiming that her mission is to wash at the water-fall the robe of her deceased lover. Although his attendants attempt to bar the maiden from the grounds of the retreat, Narukami is struck by her beauty and allows her to enter in order to hear her pathetic love story. Secretly sent to seduce him, Taema pretends to develop a sudden pain in her bosom. Narukami, to ease the supposed pain, massages her breast. Carried away with excitement, he boldly ignores priestly vows and asks her to marry him. Taema insists that they drink nuptial toasts. Unused to drink, Narukami quickly becomes befuddled. The princess prompts him to boast and draws from him the secret of how to release the Rain God. Narukami falls into a drunken stupor. Taema cuts the rope releasing the dragon Rain God and flees. Upon awakening, Narukami is so enraged that he transforms himself into a demon Thunder God. After having wrecked the monastic retreat and knocked about his terrified acolytes, he hurls himself after her in hot pursuit.

The extraordinary degree to which powerful emotions are energized in audiences confronted with the stage production of such a naive fantasy prompted us to explore what special unconscious appeal the style of Kabuki acting has. The Kabuki repertoire has a wide range, from exaggerated historical hero plays (*jidaimono*) to realistic pathetic domestic plays (*sewamono*). *Narukami* shares something of the characteristics of each.

Our exploration is grounded in certain assumptions about the nature of audiences of live theatre. What is it that provokes enjoyment and deep satisfaction for an audience interacting with live actors impersonating the characters of a fantasy?

A truism of literary and dramatic criticism has long been that a performance of a play succeeds to the extent that it induces in the audience a temporary suspension of disbelief, a momentary acceptance of the projected fantasy as if it were reality. In psychological terms, the resources of theatrical art are ultimately directed toward encouraging the members of an audience to regress. This is theatre magic.

The two main forms of Japanese traditional drama are Noh and Kabuki. The Noh play seeks to inspire wisdom and deals with the ponderable paradox, life and death. Noh is like Western classical drama with its noble simplicity and quiet grandeur. The more classical tends to be more moralistic, intellectual, ideal, abstract, with less emphasis upon the passionate and sensual feelings. This does not mean that there is no depth of feeling in the Noh or classical form; it is only of a different quality. Noh being the older is more moderate, in the sense of being less extreme in its passions, more simplified and more direct in thought, feeling, and action. In a sense it is roundabout. Noh represents the opposite pole of the younger Kabuki. Kabuki is more passionate, more detailed, more complex, more ornate, more immature, more extroverted and less introverted. Kabuki tends to permit greater activity, greater mobility and greater participation on the part of actors and audience. Noh tends to encourage greater contemplation, reflection, turning inward.

The traditional style or patterns (*kata*) of Kabuki developed out of practical trial-and-error show business over a period of four hundred years. Rich in practical showmanship and psychological insights might be the style of Western theatre today if the acting troupe of Shakespeare's Globe theatre had not been cut off in its prime, if like the Kabuki theatre of Japan it had continued uninterrupted for a period of four hundred years or more!

As one analyzes *kata,* one is led to the conviction that such a style of performing is neither accidental nor curiously or quaintly exotic. Rather it is the quintessence of a highly refined and profound feeling for what moves powerfully and deeply an audience.

What are some of the characteristics of *kata,* this high style of acting? In this mythoid tale in which the margins of what is real and what is unreal are blurred, where a priest with magical powers has succeeded in imprisoning the Rain God, priest Narukami moves slowly and delivers his lines with a paced, stentorian, orotund delivery. He is leonine (see photograph). The traditional stylized make-up with its bold use of dead white and broad contrasting black strokes lining the features sets in high relief animal characteristics. The Japanese have long been known for their aesthetic sense: their subtle and exquisite use of many colors in combinations and their dexterous manipulation of symbols which evoke natural forms but without literal realism. Make-up, costumes, wigs, all accentuate the bulk and the height of the actors who act out the fantasy. Thus these characters have melded *man* with *animal*.

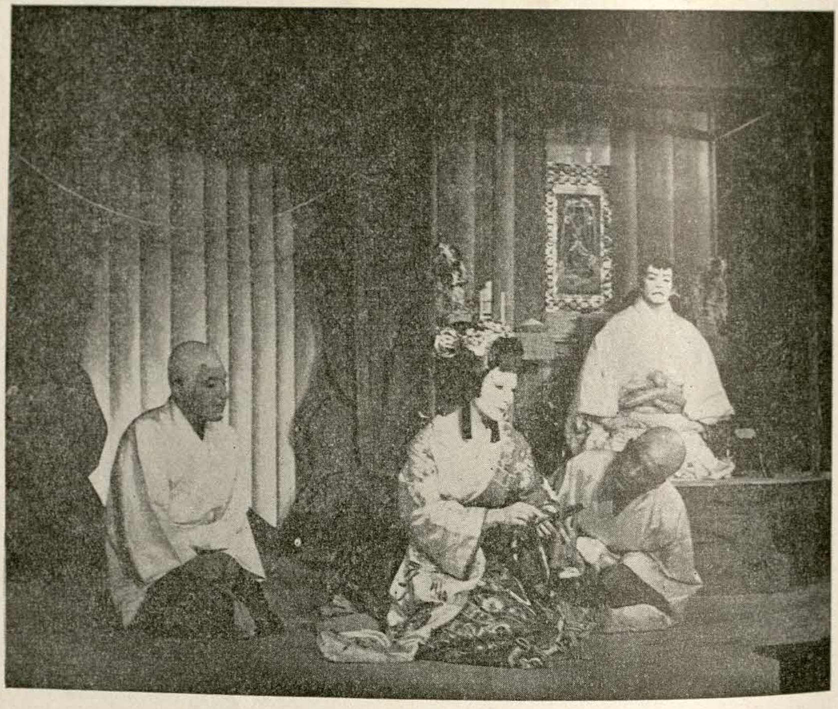


PHOTO 1—"NARUKAMI"
WIL ALBERT, JOY DILLINGHAM, MARVIN PEISNER, STEPHEN DALEY

Members of the audience are lured into vicarious participation in the fantasy unfolding before them upon the stage. The exaggerated movements and the animal characteristics of the characters are propitious for promoting associations with childhood. Its qualities help to facilitate regression. The actions and the movements are gigantic, like those of the giants that peopled our world when we were children. Giants are the parents (adults) viewed by children looking up from the floor.

If a form of theatre emphasizes the magnitude of voices, sounds that are unknown and not understood, outsized movement, grand and colorful costumes, big people, animal-likeness, big emotions, it encourages ambiguity about reality and reminiscence of times past. (See photograph #2). Childhood thinking is less realistic, and given to the recognition of human qualities only in their extreme.

Kabuki reflects folk beliefs, piety, superstition, social and didactic problems, violence and sexuality, and most of the other experiences and motifs of human life. And by a variety of mental mechanisms each onlooker identifies with one or another character, feels his feelings, shares his joy and suffering. The audience at Kabuki is neither passive nor impassive.

Sixteen generations of Kabuki actors who had chosen to build on the previous generation's practical theatre knowledge ended up leaving very little to chance in their efforts to effect a strong identification on the part of the audience with the characters of their plays. Both the music and a wide range of percussive sound effects are utilized to underscore the exaggerated and the deliberate measured movements and gestures of the actors. The cacophony of sound pulsates and moves ultimately to a series of crescendos paralleling the climax and the denouement of the story. In the case of the priest Narukami, who has lost his magical powers through the wiles of a pretty maiden, he turns himself into a demon and as a demon pursues the departing maiden out through the auditorium by means of a conventional Kabuki run-way, the *hannamichi.*

Very much a part of and blended with the animal characteristics of the Kabuki style of acting is the slightly jerky, doll-like movements of the actors. How this characteristic of the acting came about is well known. In the Eighteenth Century in Japan, the Doll Theatre, *Bunraku,* achieved sudden and great popularity with Japanese audiences. So much so that the Kabuki theatre of live actors was hard-pressed to survive in competition with the doll performers. The solution for Kabuki, which enabled it to survive, was not only to take over into its repertory the successful plays of the doll theatres, but also the actors began to incorporate into their style of acting many of the histrionic characteristics of these one-third life-sized dolls. The suggestion that the performers are not quite real but are in many ways dolls also promotes regression. In Kabuki as well as in Noh all performers are men. This polymorphous view of sexuality, where one person may play both roles, is also more characteristic of childhood.

PHOTO 2—"NARUKAMI"
JOY DILLINGHAM, STEPHEN DALEY

As was very much in evidence in observing audience reaction as well as our own to these performances of *Narukami* (in which the use of English made the play directly accessible to us as non-Japanese), we become more like children. Deep emotions began to be stirred and to make themselves felt and be acted out at a distance. For example, Narukami slowly begins to rise from his drunken stupor. During the short time in which he was lying on the floor, his wig had been changed to that of a demon wig, hair standing up and suggesting the bushy mane of a lion. With giant-like strides and resonant bellowing he begins to vent his hyperion wrath. He destroys sutras, holy scriptures (see Photograph #3); he picks up one of his acolytes and throws him through the air; he lifts and tosses a boulder (of *papier maché,* of course) which seems beyond human strength. Such giant anger and actions excite the audience and provide some increase and release of inner tensions.

Knowledge of Japanese culture supports the conclusion that there is a quality of the inevitable that the Japanese should have created this theatre. The need has been so great for a socially acceptable mechanism by means of which, at least vicariously, emotions both strong and weak can be given expression. In all areas of traditional Japanese social life the mandate of decorum is that all outward show of emotions is a violation of etiquette.

Japanese customarily carried tucked away in their kimonos a wad of paper on which they would bite when required to suppress their strong feelings of grief or anger or fear, as an aid in forcing back any outward show of emotion. For the Japanese, the theatre was and continues to be a safety valve. For the Westerner who travels to Japan and discovers and ultimately grows somewhat frustrated by the masks behind which his Japanese friends seem largely inaccessible, it comes as a great surprise to discover the displays of emotion that are indulged in by Japanese at a Kabuki performance.

At first, as a Westerner, one doubts one's perception of what is being heard and seen. Over a sentimental scene dripping with pathos on the stage, a Japanese gentleman sitting nearby is racked with loud sobs of grief; his face will be shiny with the tears which copiously flow from his eyes. Unlike any Western man at the movies or theatre, neither the Japanese man who is crying is embarrassed nor are the Japanese sitting about him. They simply take no notice, or they are as fully lost in their own loud displays of emotions, such as shouting the actor's *yago* (a type of stage name) to express passionate approval of his style and performing. No other public situation in Japan is known to the writers as being an occasion where the Japanese may with traditional Japanese acceptance, vent strong feelings. Kabuki provides them with a chance to express such feelings without limits in a socially approved environment.

If over a period of three hundred or four hundred years, this deep identification with stage fantasy and its excitation of strong emotions had

PHOTO 3—"NARUKAMI"
STEPHEN DALEY

proved either deleterious or dangerous to the state, one can be certain that the dictatorial Shoguns would have suppressed Kabuki. Unlike Noh, Kabuk was for the lower classes and may be viewed as a way in which the masses were permitted by the ruling classes to have some release for both libidinal and aggressive feelings. This brings us to an important evaluation of *kata,* the style of Kabuki acting, and its meaning for the unconscious. The aberrational violence of the fantasy being acted out, for example in a play like *Narukami,* and the attendant excitation of infantile emotions are not damaging, because the audience was and is always aware that it is a fantasy. The doll-like and the animal-like characteristics of the acting, for example, communicate to the audience, at all points, the basic unreality of what is happening.

Certainly if the members of a contemporary audience, such as attended performances of *Narukami* in New York, had begun to sense that they were being swept away by deep emotions to the point of rioting, their anxiety would have quickly inhibited the essential to effective theatre, namely, the regression. Thus the unreality factor of the patterns and style of Kabuki is ever re-assuring, and therefore, the audience with security may identify and permit itself to succumb to an accumulation of emotions that ultimately at the climax, burst like a bomb with the hero-priest catapulting himself down the run-way through the audience and out back.

We have discussed Kabuki and the ways it facilitates regression. But all drama has this quality. Is there something specific to Kabuki and the Japanese?

This may be better understood if we keep in mind that Japanese culture is accumulated from other, older sources or mother cultures. In the main, the Japanese is a derivative culture. Not unlike Rome, Japan tends to utilize ideas, institutions and techniques from other related cultures. The Japanese are accreters and coordinators. This does not deny that the Japanese have a capacity for being creative, because re-organizing the given and perhaps adding something of their own is an act of creation.

In general, the Japanese is not permitted to express feelings as easily as the Italian, Frenchman, or even American, except perhaps under the rigid puritanical controls of Colonial times. There are doubtless many reasons why the Japanese do not express their feelings, why they seem to be repressed and suppressed. But the large mass of Japanese live in tissue-paper-thin houses where everything, even loud breathing, is heard by everyone else. With extended families including several generations and many children, it was necessary in the house not to make large noises; to retain some little privacy for oneself; not to infringe upon the other person's privacy; not to intrude on the thinking, feeling and activity of other members of the family, especially with due respect for the rank ordering of the familial hierarchy. Only the children had the freedom to be immature. But even they tended to imitate the elders in the attempt to control the outward, the exaggerated, the gross expression of feeling. So the Japanese avoid the expression of extremes of inner or outer feeling but show rather the mild and moderate façade. The façade is always a controlled middle tone, whereas beneath there may be deeper feelings, and passions, thoughts and values.

This middle tone, this middle position, is the way of life for the mass of Japanese. The façade is a way of controlling for the moment. It is in a sense *kata,* a style, comparable to the Greek feeling of ecstasy. It is stepping out of one's self, a moment of hesitation before action.

We wish here to introduce another hypothesis relative to the nature of theatre experience, namely, the playwright through his fantasy is overtly or convertly seeking change. Another analysis of *Narukami,* as first conceived by Tsuuchi Hanjuro, is appropriate.

During the long closed period of Japan, which parallels the beginnings and the development of the Kabuki theatre, the samurai was the privileged class. The peasant farmers essential to Japan, as to any country, formed the broad base of the social pyramid. The merchant class was status-less and not only continued to exist, but also to grow in strength and affluence. The Popular Kabuki evolved, in part, to meet the needs of the merchants for diversion, something on which to spend their money, and to satisfy their frustrated desires for conspicuous display in clothes and furnishings.

During its long existence Kabuki has appealed especially to women. They have been among the most loyal of its audiences. As one reflects on the static society that was the Tokugawa Shogunate, it seems rather daring on the part of the playwright to introduce a character like Princess Taema who challenges the power and the authority of priest Narukami. This may explain why the character, before she climbs the rock to cut the rope to release the Rain God, prostates herself before the drunken, slumbering priest and makes an elaborate apology for the act of disobedience that she is about to perform.

Since the first audiences for the play were members of the merchant class of Japan, who were known to be restive under the repressive laws of the Shogun, *Narukami* must have been a stimulating drama: this courageous challenging of authority. What undoubtedly strengthened their identification with Princess Taema was the fact that the priest Narukami's caprice of shutting up the Rain God is reported as bringing misery to the people. Kabuki was in its long history a *popular* theatre.

Another aspect of the audience's identification with Princess Taema is the fact that she is pointedly referred to in the text of the play as coming from the Imperial Court. Princess Taema would have been impersonated by an *actor* not an actress; this, too, may have had subtle impact on audience identification. Since both Japan and the Emperor were under the ruling thumb of the Shogun, it may be assumed that over the long three hundred years, many of the land looked to the Emperor for liberation. Ultimately it did come and erupted in the 1860's, a movement to become known as the Meiji Restoration.

However, identification would not have been exclusively with the character Princess Taema. As in a dream, the audience has shifting identifications or may identify with several or all of the characters of the fantasy. Undoubtedly, the repressed merchant class audience would have identified, strongly and with deep satisfaction, with Narukami as he converts himself into a demon Thunder God. As such, he acts on his ire indiscriminately. He turns on his own acolytes. He destroys holy scriptures. Of course, all of this activity is formalized and exploited for aesthetic appeal. As we noted for a contemporary audience, so for the first audiences of Japanese merchants. By holding the complete acting out of destructive and aggressive feelings within aesthetic and traditional patterns (*kata*), Kabuki achieved then as well as now, exceptional catharsis. It does not pass beyond a threshold of tolerance for an inter-reacting audience.

When a person comes out of the Noh playhouse, he is supposed to come out wiser. When he comes out of the Kabuki-za, he comes out entertained, refreshed, catharted. He has rid himself of something rather than having gotten something. He comes out as if he had been in an hypnotic trance, in a dream. It has been like a dream which permits us to sleep and awaken again, refreshed, recuperated and ready to face a reality much less acceptable and satisfying than the Kabuki story.

REFERENCES

1. BENEDICT, R. The Chrysanthemum and the Sword. New York: Houghton Mifflin Company, 1946.
2. BOWER, F. The Japanese Theatre. New York: Hermitage House, 1952.
3. ERNST, E. Kabuki Theatre. New York: Oxford University Press, Inc., 1956.
4. FREUD, A. The Ego and the Mechanisms of Defense. New York: International Universities Press, 1946.
5. FREUD, S. The Relation of the Poet to Day-Dreaming (1908). Collected Papers, Vol. IV. London: Hogarth Press, 1949.
6. HALFORD, A. S. & G. M. The Kabuki Handbook. Tokyo: Charles E. Tuttle Company, 1956.
7. KIRKUP, J. Japanese impressions. *Today's Japan,* 1960, **5** (10), 81-88.

8. MITCHELL, J. D. Andre Gide, rebel and conformist. *The American Imago,* Summer, 1959, **16** (2), 147-153.
9. ———. Applied psycho-analysis in the director-actor relationship. *The American Imago,* 1956, **13** (3), 223-239.
10. ———. Applied psychoanalysis in the drama. *The American Imago,* 1957, **14** (3), 263-280.
11. ———. Contemporary American Theatre . . . a psychologic sounding board. *Natya, Theatre Arts Journal,* Spring, 1960, **4** (1), 1-10.
12. ———. A Sanskrit Classic: Shakuntala. *In* de Bary, W. T. (*Ed.*) Approaches to the Oriental Classics. New York: Columbia University Press, 1959, 119-131.
13. ———. The Sanskrit Drama Shakuntala: A psychologic sounding board for Hindu culture. *The American Imago,* Winter, 1959, **16** (4), 229-348.
14. ———. The Theater of India and Southeast Asia. *In* Asia and the Humanities. Bloomington, Indiana: Comparative Literature Committee, Indiana University, 146-155.
15. MITCHELL, J. D., & DREW, G. Ein Besuch Im Neuen Kabuki-Theater in Osaka. *Buhnentechnische Rundschau,* June, 1960.
16. MITCHELL, J. D., DREW, G., & MITCHELL, M. P. The Moscow art theatre in rehearsal. *Educational Theatre Journal,* December, 1960, *XII* (4) 276-284.
17. SCHWARTZ, E. K. A psychoanalytic study of the fairy tale. *Amer. J. of Psychother.,* 1956, **10**, 740-762.
18. ———. Psychotherapy in Mexico. *In* Masserman, J. H., & Moreno, J. L. (*Eds.*) *Progr. in Psychother.* New York: Grune & Stratton, 1957, 216-223.
19. ———. Some points of view on dreams and dreaming. *J. of Long Island Consult. Ctr.,* 1961, **1** (3), 9-14.
20. SCHWARTZ, E. K., & WOLF, E. K. Irrational trends in contemporary psychotherapy: Cultural Correlates. *Psychoanal. & Psychoanal. Rev.,* Spring-Summer, 1958, 65-74.
21. SCOTT, A. C. The Kabuki Theatre of Japan. London: George Allen & Unwin, Ltd., 1955.
22. Special Issue on Schizophrenia. *Rev. and Newsl. Transcult. Res. in Ment. Health Prob., December,* 1960, **9**.
23. YONEZO, H. Kabuki. Tokyo: The Kenkyusha Press, 1956.

Institute for Advanced Studies
in the Theatre Arts
418 W. 42nd Street
New York 36, New York

Postgraduate Center for
Psychotherapy
218 East 70th Street
New York 21, N.Y.

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 281-285.

PSYCHOSOMATIC V*[1]

Veterans Administration Mental Hygiene Clinic, Oakland, California

HERBERT C. ARCHIBALD, DOROTHY BELL, CHRISTINE MILLER, AND CLARE W. THOMPSON

A. INTRODUCTION

The time-honored approach in personality test research has been to administer tests to individuals grouped by psychiatric diagnoses in order to discover which test scores or patterns serve to differentiate these groups. An opposite approach involves first selecting specific scores or patterns which characterize the performance of some persons on the test. An attempt can then be made to determine how individuals who obtain these scores or patterns differ from the remainder of the population which was tested. The characterization of persons selected on this basis need not be restricted to indicating their diagnostic status but could include other psychiatric data and demographic data as well.

We have sought to determine what reasonable inferences can be made about patients whose Minnesota Multiphasic Personality Inventory (MMPI) profiles show the presence of the psychosomatic V (1). This configuration is a frequently observed pattern of scores which we believe to convey a considerable amount of clinical information. For the present study, we have defined the psychosomatic V as a configuration which exists when T-scores for the Hypochondriasis (K corrected) and the Hysteria scales exceed 70, and the T score for the Depression scale is more than 5 T-score points below the lowest of the two former scales.

B. METHOD

The population selected for study in the present investigation consisted of male patients of the Veterans Administration Mental Hygiene Clinic, Oakland, California. These were veterans of World War II and Korea with service-connected neuropsychiatric disorders. As part of the routine intake procedure, all were asked to take the MMPI. The experimental

---

*Received in the Editorial Office on May 19, 1961 and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] From the Veterans Administration Regional Office, San Francisco.

group consisted of those 66 patients with a psychosomatic V. The control group data were obtained from a continuing survey of the total clinic population, carried on from 1952 to 1958 (2). The experimental and control groups were compared for a number of variables available in clinic records.

C. Results

1. *Incidence of the Psychosomatic V*

Sixty-six profiles with this configuration were found in the total pool of 784 MMPI's administered to patients at the clinic, an incidence of 8.7 per cent. Among a group of 59 veterans who declined treatment, six such profiles, constituting 10.2 of the sample were found. A chi square test of the significance of the difference in the incidence of the psychosomatic V in the two groups yielded a value of .18 ($P > .05$). Consequently, there is no evidence that this profile is either more or less characteristic of individuals seeking psychiatric treatment than of those who decline such treatment.

No significant differences between the experimental and control groups were found for age, race, education or number of hospitalizations. Nor were significant differences found for sex of therapist or discipline of therapist. The experimental group were, however, more frequently engaged in white collar occupations and also less often unemployed than was true for the control group ($P < .02$). It should be mentioned that although there was not a significant difference between the control and experimental groups in education, the average educational level is slightly higher for the control group. Despite a slight handicap over the control group in education, the experimental group are more often employed and at a higher lever than the controls.

The diagnoses given the psychosomatic group were compared with those given to the control group. The diagnoses were grouped into four categories: Anxiety Reaction, Psychophysiological Reaction, Psychotic, and Other. It is of interest that significantly more patients with a psychosomatic V are diagnosed as "Psychophysiological Reaction" ($P < .01$). However it is perhaps of more interest that there were significantly fewer psychotic diagnoses in the psychosomatic V group since it will be recalled that in the selection of cases, no attention was paid to the elevation of any MMPI scale other than the first three. For example, it was not necessary for the Hypochondriasis and Hysteria scales to be the two highest scores. Yet significantly fewer of these cases than of the total clinic population were diagnosed as psychotic by the examining psychiatrist ($P < .02$). Moreover, although it was specified that the Depression scale, primarily a measure of anxiety, be at least five scale points lower than Hypochondriasis and Hysteria,

TABLE 1
COMPARISON OF GROUP SHOWING PSYCHOSOMATIC V WITH TOTAL CLINIC POPULATION

Demographic Variables		Experimental group	Control group*	χ^2	P
Race	White	56	304	0.33	0.60
	Negro	9	43		
Age		M=31.8 yrs.	M=31.8**	3.85	0.30
Marital status	Single	7	70	4.25	0.20
	Divorced	5	35		
	Married	54	247		
Occupation	White Collar	33	110	6.51	0.02
	Blue Collar	33	227		
Employment status at intake	Employed	42	189	6.18	0.02
	Unemployed	13	119		
	In Training	10	31		
Education		M=11.7 gds	M=12.0 gds	5.07	0.10
Hospitalizations		M= 0.93	M= 0.94	2.72	0.50
Therapy variables					
Sex of therapist	Female	18	123	0.16	0.70
	Male	48	215		
	Psychiatrist	22	91	2.94	0.30
	PSW	16	70		
	Psychologist	26	179		

* The variables shown in the table varied slightly from year to year; consequently the size of the control group varies depending upon the particular measure under consideration. The same is true of the experimental group. Nevertheless, many meaningful comparisons can be made between the experimental group and the clinic population as a whole. Unless otherwise specified, this is the '54 to '56 group, with an N varying with the measure from 343 to 410.

** N is based on the '52 to '53 group with an N of 130. χ^2 is based on the usual '54 to '56 group.

nevertheless as many of these patients were diagnosed "Anxiety Reaction" as in the total clinic population. Most of the latter patients did show evidence of muscular tension on examination.

These differences were further corroborated by the presenting complaints of the experimental group and of a control group of 66 subjects composed of every third patient seen in this clinic during September 1957. The symptom list was issued by the VA Central Office Department of Medicine & Surgery in connection with a study of mental hygiene clinic treatment. It consists of 67 common presenting symptoms and complaints. The mean number of symptoms was significantly lower in the experimental group (mean for experimental group = 5.2, $SD = .28$, mean for control group is 8.7, $SD = 2.3$).

Those which were more characteristic of the experimental group ($p < .05$) were: back pain, dizziness, feeling of weakness, forgetfulness, headaches, nausea, nightmares, stomach pains, tremors, vomiting, weight loss. The more characteristic symptoms for the control group ($p < .05$) were: apathy, apprehension, bizarre thinking, compulsive behavior, concentration difficulties, depression, dislike of crowds, ideas of reference, jitteriness, obsessive thinking, passivity, restlessness, restless sleep, seclusiveness, tension, withdrawal. It will be noted that the symptoms of the experimental group contrasted with those of the control group in showing more somatic complaints and fewer psychotic symptoms.

It is of interest to see how long these patients with primarily somatic complaints remain in psychotherapy. A comparison of the psychosomatic V group with the controls on number of interviews shows that the psychosomatic V group tended to remain 10-19 interviews significantly more frequently than the controls ($P < .01$). There were no significant differences for the 0-1, 2-9, or more than 20 interview categories.

It has been found at this clinic that therapists' ratings of patient improvement are highly correlated with patients' length of stay, and a very few "improved" ratings are given patients who stay less than ten interviews. Almost one-half of the experimental group came for at least 10 interviews. It is of interest that more of the psychosomatic V group came for at least 10 interviews than was true of the "Psychophysiological Reaction" and "Conversion Reaction" groups whom they might be thought to resemble. Perhaps of more importance than length of stay is the finding that the experimental group were more often rated improved than the control group (chi square 25.4 for 2 degrees of freedom, $P < .001$).

D. DISCUSSION

It is the authors' impression that patients with a psychosomatic V have a poor reputation with regard to prognosis in psychotherapy. Because this configuration is interpreted to indicate denial of anxiety, focus on physical symptoms and a naive uninsightfulness about psychological functioning, they are thought to be poor candidates for psychotherapy. Our data, however, do not confirm the validity of this clinical folklore since the psychosomatic V patients stayed longer in therapy and were rated improved more often than the control group. Despite a larger number of physical symptoms and a slight handicap in education, they were more frequently employed at a higher occupational level than other clinic patients. This suggests a conventional, integrated approach to life that is reflected in their ability to remain in psychotherapy and profit from it.

E. SUMMARY

Sixty-six patients with psychosomatic V patterns on the MMPI were compared with the general clinic population on several non-test variables. Patients with the psychosomatic V were more apt to be employed, were more often engaged in "white collar" as opposed to "blue collar" occupations, and complained or exhibited more physical symptoms than the general population at the time of intake. After becoming patients, the psychosomatic V group tended to stay in therapy longer and to be rated improved more often than the general clinic control group.

REFERENCES

1. HATHAWAY, S. R., & MCKINLEY, J. C. The Minnesota Multiphasic Personality Inventory Manual (Revised) New York: Psychological Corp., 1951.
2. SULLIVAN, P., MILLER, C., & SMELSER, W. Factors in length of stay and progress in psychotherapy. *J. Consult. Phychol.,* 1958, **22**, No. 1, 1-9.

VA Mental Hygiene Clinic
1305 Franklin Street
Oakland 12, California

Published as a separate and in *The Journal of Psychology*, 1961, 52, 287-301.

CONVERGENCE AS A CUE TO ABSOLUTE DISTANCE*

U.S. Army Medical Research Laboratory, Fort Knox, Kentucky

WALTER C. GOGEL[1]

A. PROBLEM

When fixating an object the eyes are positioned (converged) so that light from the object falls on the fovea of each eye. It is possible that the activity of proprioceptive receptors in the eye muscles involved in this convergence might provide a cue to the distance of the object fixated. Summaries of research relevant to this problem are given by Baird (1), Osgood (9), Woodworth (11), and Woodworth and Schlosberg (12). But in spite of the amount of attention it has received this issue has not been decisively settled. One reason for this is found in the difficulty of isolating the perceptual effects of accommodation and convergence from each other and from other possible distance cues. For example, if the accommodative-convergence reflex is used to induce convergence from monocular stimulation, then accommodation as well as convergence is present as a distance cue. On the other hand, binocular stimulation for various convergences can be used with accommodation held constant. In this case, the binocular (diplopic) images are themselves a possible cue to distance and also the constant value of accommodation would tend to negate the perceptual effects of the convergence changes. It is apparent from the studies in this area that the problem of isolating the perceptual effects of either accommodation or convergence is difficult. The present study used a situation in which a very limited number of cues were available to determine the apparent distance of the binocular object. Upon determining whether a perception of distance occurred as a function of these cues, the number of available cues was changed in an effort to find the limiting conditions under which the perception of the distance of the binocular object would occur.

For the purposes of this study, a distinction must be made between a perceived absolute and a perceived relative distance and also between an absolute convergence and a convergence difference. A perception by the observer that an object is a definite particular distance from himself is a perception of absolute distance. A perception of a depth between objects, or between different distance positions of the same object at different times, is a perception of relative distance. Absolute convergence is the convergence of the eyes at a particular value, while a convergence difference is a change in convergence from one value to another. In the present study, the experimental question is whether absolute convergence is a determiner of perceived absolute distance.

---

* Received in the Editorial Office on June 5, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The author wishes to thank James M. H. Gregg and Charles A. Wainwright, Jr. for their help in collecting and analyzing the data of this study.

Perceived absolute distance can be measured in several ways. Bourbon (2), for example, measured the perceived absolute distance of a circular spot of light by means of distance estimates. Using relatively far distances, he found that convergence was inadequate as an indicator of the perceived distance of the spot of light. Baird (1) presented a straight edge at distances from 29 to 90 cm. He found that distance estimates of this object varied directly with the physical distance with the usual error being one of underestimation. While the measurement of perceived absolute distance by means of verbal estimates is meaningful, it has the disadvantage that the measuring standard must be supplied by memory rather than being directly available. More direct methods of measurement have been attempted. In a study by Swensen (10) the *S* moved a non-visible object with his hand to the same apparent distance as the stimulus object. The convergence and accommodation to the stimulus object was for either 25, 30, or 40 cm. and with all combinations of these convergence and accommodation values being used. Swensen found good discrimination between these distances with accommodation being less effective than convergence. Carr reports (3, p. 251) that Duncan, using the same apparatus, obtained results which indicated that perceived distance varied directly with the convergence and accommodative distance. Grant (6) like Swensen used a haploscope but with an improved control over the accommodation to the binocular object. As a measure of perceived absolute distance, the *S* adjusted a monocularly observed penny to the perceived distance of the binocular object. The accommodation and convergence to the binocular object was for 25, 33, and 50 cm. with the two cues sometimes in agreement and sometimes in opposition. The results suggest that some, but not a large, amount of discrimination occurred between the three distances. This last method used the perceived distance of the familiar object (the penny) as a measure of the perceived distance of the binocular object. However, Gogel, Hartman and Harker (5) have found that the retinal size of a familiar object is not an adequate cue to perceived absolute distance. It is possible therefore that perceived relative rather than perceived absolute distance was being measured as a consequence of the successive presentations of the binocular object in the study by Grant.

The problem of the possible relation between convergence and perceived absolute distance is of some interest for its own sake. Probably more important, however, is its theoretical consequences. If convergence is a sensitive determiner of perceived absolute distance, it may be a significant factor in determining other binocular phenomena. One such phenomenon is the perceived depth between two binocular objects which occurs even though only binocular cues are present. It seems that except for the study by Bourbon (2), the possible relation between convergence and perceived absolute distance has not been investigated for distances beyond approximately one meter. Approximately 4 degrees of convergence change remains between 90 cm. and infinity. If the perceptual consequences of convergence are so imprecise as to not permit absolute distance discriminations within this range of convergence values, convergence cannot have theoretical significance for the other binocular perceptual phenomena. Also, it is possible that the threshold for detecting a change of perceived absolute distance depends upon the magnitude of the absolute convergence used. Physical linear distance changes more rapidly with convergence at the small than at the large convergence values. Thus, it might be expected that an equal change in convergence at small convergence values would result in a greater perceptual change than at large convergence values. If so, the greatest change in perceived absolute distance per unit of convergence change would be expected from situations involving small values of absolute convergence. In the present study, the range of convergence values was from 0 to 4 degrees.

B. EXPERIMENTS I AND II

1. *Apparatus*

a. *The Visual Alley*. Since large distances from *S* were often used in this study, a method of measuring quite large perceived absolute distances was required. For this purpose, *S* was provided with a visual ruler which consisted of a monocularly observed alley, 34.2 feet long and 4.2 feet wide. This alley contained a number of distance cues, for example those of perspective and size, such that the alley appeared to *S* to extend away from himself in distance. This monocularly viewed alley constituted a measuring stick against which the distance of the binocular object could be judged.[2] This insured that to whatever extent absolute distance was perceived with the binocular object, the situation for its precise measurement was simultaneously available.

---

[2] The use of a comparison field for the measurement of apparent distance has been discussed by Ittelson (7) and by Kilpatrick (8) in conjunction with some of the Ames demonstrations and related experiments.

A schematic diagram of the monocularly viewed alley is shown in Figure 1. The floor of the alley was covered with a checkerboard pattern of light and dark grey rectangles. White vertical cardboard rectangles were distributed uniformly in distance along the alley. These cards were numbered consecutively with increasing distance, with the nearest card numbered one and the farthest numbered fifteen. The nearest white card was 5.3 feet from *S* with 2.1 feet between cards. As is indicated in Figure 1, the numbered cards were also distributed right and left so that no card was obscured by another.

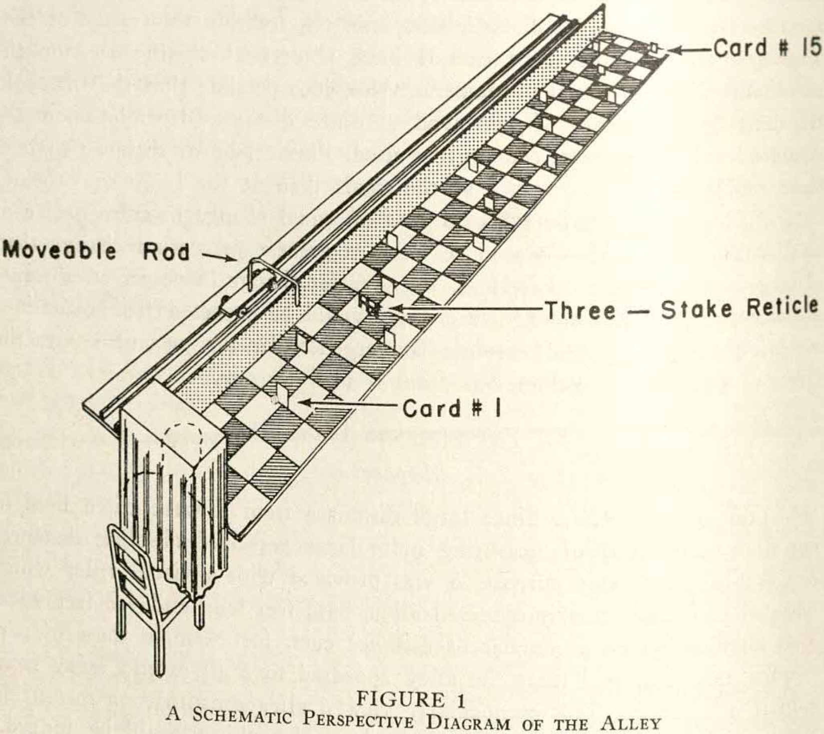


FIGURE 1
A SCHEMATIC PERSPECTIVE DIAGRAM OF THE ALLEY

By means of a hidden cart shown on the left in Figure 1, *S* could adjust a rod to the apparent distance of any of the numbered cards. This adjustment was a control to determine that distance was perceived in the monocular visual alley.

b. *The Stake Reticles.* The binocular object which was presented at the different convergence values is shown in Figure 1 as an object composed of three vertical lines connected by a horizontal line. This object called the three-stake reticle was used throughout Experiments I and II. The stakes of the three-stake reticle were each 7 minutes wide and 20 minutes high with 58 minutes between stakes. In Experiment II, small monocular components were added to the three-stake reticles in order to be certain that neither of the two images comprising the binocular reticle was suppressed when binocular fusion was reported.

The three-stake reticle was not a real object actually located in front of *S*. Instead it was stereoscopically generated by using a pair of reflex sights with one reflex sight before each eye of *S*. Each reflex sight contained an aperture source of light of the shape required for the three-stake reticle. The light from this source passed through an optical system and emerged as parallel light. This parallel light was then reflected from a reflecting-transmitting surface into an eye of *S*. The principles involved in this instrument are discussed in a previous study involving the generation of a stake reticle (4).

In the stereoscopic instrument, the reflecting-transmitting surface in front of the right eye permitted *S* to see both the alley and the reticle with his right eye. However, a black opaque cover behind the left reflecting-transmitting surface permitted *S* to see only the reticle image with his left eye. Neither the alley nor any other object except the reticle was visible with the left eye. This means that except for the stake stimulus, the left eye was in darkness.

By turning an instrument knob, *E* could rotate the reflex sights in opposite directions and by this means, could present to *S* the various values of convergence to the stake reticle required in the experiments. Throughout this study, the convergence to, or the convergence value of, an object or reticle is the angle between the visual axes of the two eyes which results when the eyes are positioned so that the image of the object or reticle in each eye is on corresponding foveal points.

The accommodation value of the binocular reticle was always for optical infinity regardless of its convergence value. The orientation of the stereoscopic instrument was adjusted by *E* so that the directional position of the reticle with respect to the visual alley always was constant for the *S*'s right eye for all values of convergence. The vertical position of the reticle was such that it was always higher than the tops of the cards in the visual alley. When a small black curtain was lowered, *S*'s view of the alley was cut off without disturbing the view of the reticle in either eye.

2. *Subjects*

The same twelve *S*s were used in Experiments I and II. All *S*s were men with a visual acuity of at least 20/30 in each eye, corrected if necessary, and at least 85 seconds of arc of stereoscopic acuity as measured by the Diagnostic Series in the Keystone Telebinocular. The majority of these *S*s had been used previously in one or more depth experiments involving the instrument for generating the stake reticle. Thus, while this previous experience did not involve the particular convergence task of the present study, many of the *S*s cannot be considered as having been experimentally naive. The *S*s A, D, and I were by far the most experienced of the group in the observation of binocular phenomena under laboratory conditions. The same letter in the several figures of this paper identifies the same *S* in the several experiments.

3. *Procedure*

To *S* the binocular object appeared to be suspended above the numbered cards in the alley. The problem was to determine where in distance the binocular object appeared to be as a function of its convergence value. For this purpose *S* judged the apparent distance position of the binocular object relative to the numbered white (monocular) cards for each of the several values of convergence. Six values of convergence were used varying from 0 to 4 degrees in steps of 48 minutes. These six values were presented three times to each *S* with the resulting 18 values presented in random order.

The general procedure was as follows. The *S* was seated in the viewing position and a black velveteen cloth was adjusted over his head and shoulders. This black cloth plus the removal of all extraneous light sources in the room insured that the left eye piece was not illuminated by stray light, i.e., no stimulus was ever present to the left eye except the left binocular image. The binocular object was adjusted by *E* to the desired convergence and orientation. The light source for the binocular object was turned on with the small black curtain lowered to prevent *S* from seeing the visual alley. It was explained to *S* what he might expect if he saw diplopic images of the binocular reticle. He was informed, for example, that incorrect fusion of the binocular images could result in more than three stakes being seen in the three-stake reticle. When using the three-stake reticle, the *S* was questioned frequently as to the number of stakes present. When *S* indicated that the binocular object was properly fused, the small black curtain was raised revealing the monocularly viewed alley with the binocular object continuing to be present. The *S* was asked to keep his fixation in the depth vicinity of the binocular object. The *S* then judged the distance position of the binocular object relative to the numbered (monocular) white cards. Following this, the front curtain was lowered, the binocular object was turned off, and the next convergence value was adjusted by *E*. The process was repeated until *S* had completed the required judgments.

In order to be certain that depth was perceived in the monocularly viewed alley, two control judgments were used in Experiment I. These control judgments were completely monocular. These consisted of *S* adjusting the moveable rod to apparent equidistance with cards 2, 4, 6, 8, and 10 in random order and also estimating the distance of each of these cards from himself in feet.

Two additional kinds of control judgments were used also. One of these was used in both Experiments I and II following the 18 presentations of the binocular reticle. In this control, the procedure for presenting the six convergence values of the stake reticle was repeated exactly as before except that just before the small curtain was raised revealing the monocular alley, an occluder on the left eye system was inserted to eliminate *S*'s view of the reticle with his left eye. Under these conditions, i.e., with the stake reticle as well as the alley viewed monocularly, *S* again judged the distance of the stake reticle relative to the numbered cards. This was repeated six times, once for each of the six values of convergence used previously. The purpose of this control was to determine whether binocular retinal cues were influencing the judgments. This control will be called the monocular reticle control. The process of fusing the diplopic images was completed before the occluder was inserted. The insertion of the occluder removed the convergence cue to the distance of the reticle. Therefore, the distance discriminations with the reticle would be eliminated by using the monocular reticle control only if convergence rather than diplopic images had been the effective distance cue.

The second additional control was used only in Experiment II. In this control the three-stake reticle was binocularly presented at 2 degrees of arc convergence. Then, while *S* viewed the reticle, *E* adjusted its convergence value to either 1 or 3 degrees at the rate of approximately 12 minutes of arc per second of time. The *S* reported the distance position of the reticle with respect to the numbered cards before and after the movement and also the direction of the movement. The perception of movement in the proper direction indicated that good binocular fusion was occurring under the conditions of these experiments.

4. *Results*

The average results from each of the twelve *S*s in Experiments I and II are given in Figure 2. The *S*s were permitted to extrapolate the scale of numbered cards at either end of the alley if necessary. Therefore, the reported card numbers were occasionally greater than 15 and occasionally less than 1. In general, there was little or no decrease in the reported distance of the stake reticle with an increase in convergence up to 4 degrees except perhaps for *S*s A, D, and I.

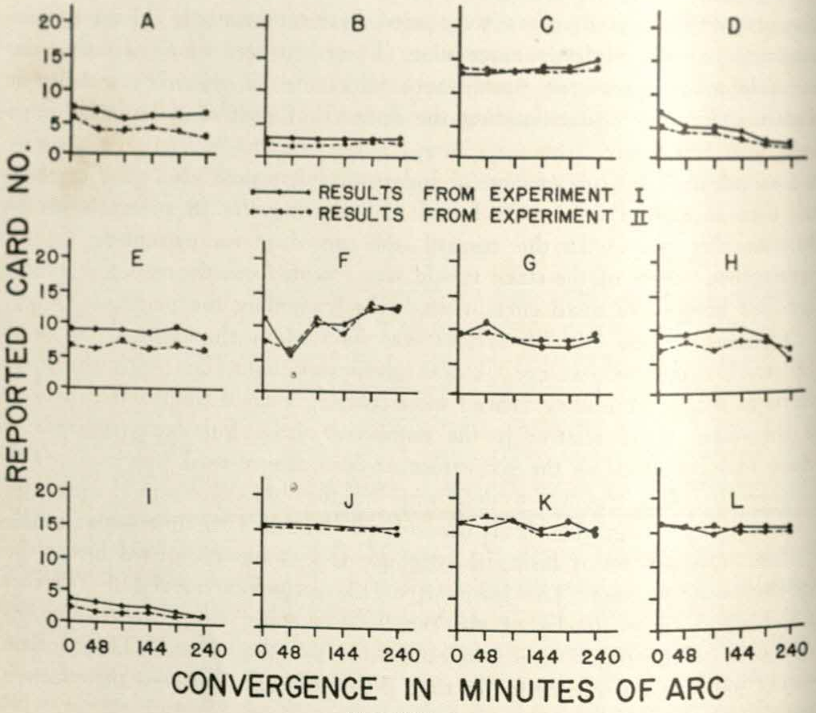


FIGURE 2

Reported Distance Position (card number) of an Object as a Function of Convergence. A Larger Reported Card Number Indicates that the Object was Perceived to be at a Greater Distance.

There is no doubt that the monocularly viewed alley appeared to extend in distance from *S* and that it constituted therefore an adequate ruler for measuring the apparent distance of the binocular object. This is indicated by the results from both the verbal reports of distance and the adjustments of the moveable rod. The magnitude of both measures linearly increased with physical distance. It is clear that the larger the card number (physical distance) the greater was the apparent distance of the card.

The results from the monocular reticle control are particularly pertinent for the *S*s whose curves in Figure 2 were not horizontal. These are *S*s A, D, F, and I. Each of these gave some indication of a decrease in perceived distance with an increase in convergence, except F who showed an opposite tendency. If convergence rather than the retinal factor of the occurrence of diplopic images is responsible for the change in perceived distance in Figure 2, for these four *S*s, then the monocular control and binocular results should differ. This was the case with A, D, and I. With F, however, the monocular control and the binocular results were similar. Therefore, the binocular results from F are not considered to be meaningful for the purposes of this study. In the case of A, D, and I, it seems that convergence, rather than the process of fusing diplopic images, provided the cue for the decrease in perceived distance with an increase in convergence.

The results from the moving reticle control indicate that only three *S*s (B, C, L) failed to make distance discriminations when continuously observing the changing convergence. Subject F gave erratic reports. The average reports from the remaining *S*s indicated that the stake reticle appeared to move 1.2 card units for the 2 degree change in convergence.

C. EXPERIMENT III

Approximately one minute of time occurred between the different presentations of the binocular stake reticle in Experiments I and II. Three *S*s (A, D, and I) in these experiments evidenced some ability to perceive the reticle as being at greater distances with lesser convergences. The problem occurs whether the change in distance judgments on successive presentations with these three *S*s was the product of the different absolute convergences or the result of the convergence differences between presentations. In the second case, the distance judgments on successive presentations would not be independent of each other. To investigate this problem, in Experiment III, the time between successive presentations was increased to 20 minutes using *S*s A, D, and I. The procedure and apparatus of Experiment III was the same as that of Experiment II except that neither of the control situations of Experiment II was used.

The average results from Experiment III compared with the average results from Experiments I and II are shown in Figure 3. Again, as shown by the slopes of the lines in Figure 3, a change in reported card number as a consequence of a change in convergence was present for these three *Ss*. It seems that increasing the time between judgments from 1 to 20 minutes did not decrease the sensitivity of the distance discriminations with these *Ss*. Thus, the evidence is that successive distance judgments were independent of each other. The distance discriminations which occurred with these three *Ss*, therefore, cannot be attributed to the method of successive presentation.

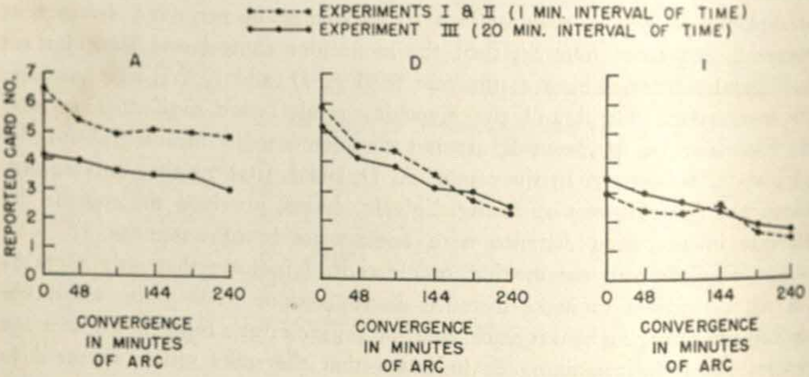


FIGURE 3
RESULTS FROM CHANGING THE TIME INTERVAL BETWEEN CONVERGENCE JUDGMENTS FOR THREE EXPERIENCED SUBJECTS. A LARGER REPORTED CARD NUMBER INDICATES THAT THE OBJECT WAS PERCEIVED TO BE AT A GREATER DISTANCE.

D. EXPERIMENT IV.

The binocular object which was generated by the reticles in Experiments I, II, and III was unreal in the sense that its value of accommodation was always to optical infinity regardless of its convergence value. In the case of a real binocular object, accommodation and convergence are in agreement for all values of convergence. It is therefore possible that the constant accommodation to the stake reticles in Experiments I, II, and III acted to detract from the effectiveness of the convergence as a determiner of perceived absolute distance. To test this possibility, in Experiment IV the perceived distance to a real object was compared with that to a single-stake reticle as a function of convergence.

The single-stake reticle was a binocular stereoscopic object generated in the same manner as the three-stake reticle of Experiments I, II, and III. It consisted of a single luminous binocular stake 7.6 minutes wide and 45 minutes high.

The real object in Experiment IV consisted of a rectangular-shaped aperture cut into black opaque cardboard and transilluminated with diffused white light. The source of diffused white light was a box containing fluorescent tubes and a translucent Plexiglas front upon which the cardboard was mounted. Four such light boxes and apertures were used and were positioned at four distances so as to duplicate four convergence values to the single-stake reticle. At these distances, the sizes of the apertures were such that the retinal subtense of the real object for each *S* was always constant and always the same as that of the previously or subsequently used single-stake reticle. The real object and single-stake reticle were identical in angular size, convergence, apparent brightness, and color. However, unlike the stake reticle, the real object presented the cue of accommodation to assist *S* in localizing it in the visual field.

The light boxes for producing the real binocular object were located along the left side of the monocular alley in an otherwise dark portion of the experimental room. Mirrors were carefully positioned so that the real object was binocularly presented to *S* and like the stake reticle appeared to be located directly in front of *S* in the monocular.alley. As in the case of the reticles, the real binocular stake was the only object visible to the left eye while both the real stake and the monocular alley were visible to the right eye. The physical limitations of the positioning of the real object were such that the largest convergence value possible in Experiment IV was less than that used in the other experiments.

In Experiment IV, four values of convergence were used varying from 24 minutes to 2 degrees in steps of 32 minutes. These four values were presented four times to each *S* (for both the real and stereoscopic object) with the resulting set of 16 values presented in random order. The same general procedure was used as in the previous experiments. All the *S*s had a visual acuity of at least 20/30 in each eye corrected if necessary and a stereoscopic acuity of at least 85 seconds of arc as measured with the Diagnostic Series of the Keystone Telebinocular.

The results from using the varying convergence with the constant accommodation can be compared with the results from varying both of these cue systems together. The average results from the individual *S*s are shown in Figure 4. It will be noted that with half the *S*s (M, N, O, T, U, W) the addition of accommodation changes appropriate to the convergence changes increased the ability to discriminate distance. With the other six *S*s the addition of accommodation produced no perceptual effect.

In still another experiment with the 12 *S*s of Experiment IV, the real stake was presented monocularly. This resulted in an average curve which was very similar to that from binocular observation of the real stake. In this last experiment, both accommodation and accommodative convergence were cues to apparent distance but the possible distance cues involved in the fusion of diplopic images was absent. It seems that the fusion of diplopic images was not responsible for the apparent absolute distance localization which occurred previously in this study with some *S*s.

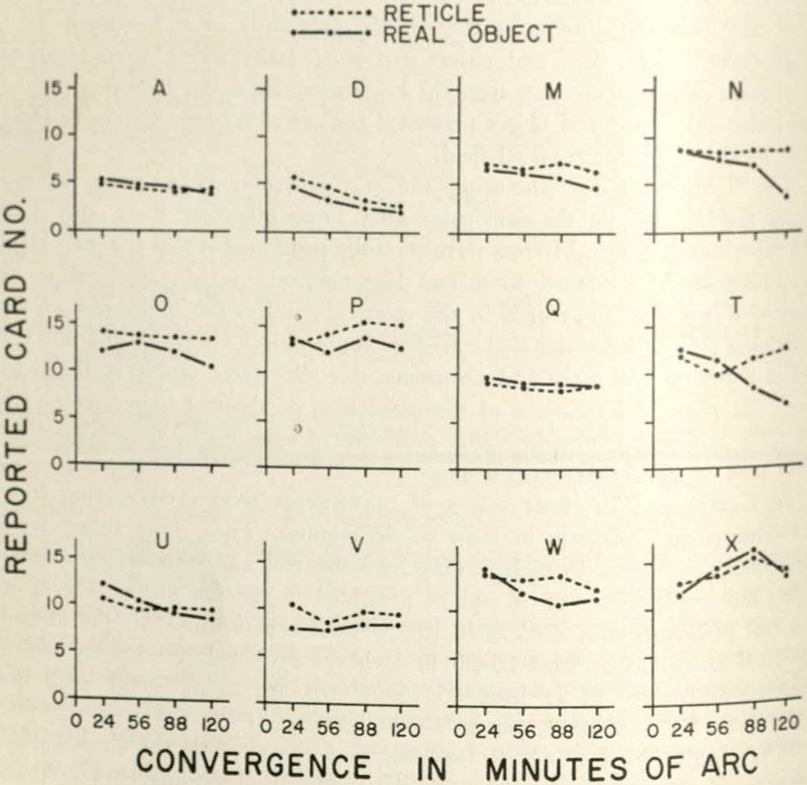


FIGURE 4

THE RELATION BETWEEN CONVERGENCE AND REPORTED DISTANCE WHEN USING A STEREOSCOPICALLY GENERATED BINOCULAR OBJECT (DOTTED LINE) OR A REAL BINOCULAR OBJECT (SOLID LINE). A LARGER REPORTED CARD NUMBER INDICATES THAT THE OBJECT WAS PERCEIVED TO BE AT A GREATER DISTANCE.

E. Discussion

As was mentioned previously, two kinds of convergence should be distinguished. Absolute convergence refers to the particular steady value of convergence resulting from a particular stimulus. In these experiments this was controlled by the fusion of the binocularly presented object. This fusion was achieved prior to lifting the small curtain which permitted *S* to see the monocular alley. A convergence difference refers to a change of convergence. This was not likely to occur in this study during a single presentation of a particular convergence value since, for any particular presentation, the convergence was fixed by the fusion of the binocular stimulus. Also, the lack of consequences resulting from changing the interval between presentations from 1 to 20 minutes reduces the possibility that the convergence differences between successive presentations were a distance cue. Therefore, it is unlikely that the changes in perceived distance for *S*s A, D, and I in Experiments I, II, and III can be explained by any factor other than that of absolute convergence. However, these changes in perceived distance were not large. For example, the *S* who showed the largest changes as a function of convergence only reported an average change of approximately 8 feet for a convergence change physically equivalent to a change from infinity to about 3 feet. As indicated by successive experiments, however, these changes though not large were reliable. It is concluded that in Experiments I, II and III absolute convergence with a few *S*s was a determiner—though not a very sensitive determiner—of perceived absolute distance. The three *S*s who reported a change in perceived absolute distance with changes in absolute convergence were, of all the *S*s, most experienced in viewing binocular phenomena under laboratory conditions. It is possible therefore that with special training other *S*s would have shown similar behavior. Apparently, however, ordinary visual experience was not sufficient to produce this type of discrimination.

Two kinds of accommodation also should be distinguished. Absolute accommodation reflects the accommodative value of a particular object by itself. The absolute accommodation to the binocular object in this study was held constant for all convergences when the reticles were used and was changed with convergence when the real object was used. Accommodative differences reflect the differences in accommodation between objects. When the reticles were the binocular object, both the constant value of the absolute accommodation to the reticle and the lack of accommodative differences between the reticle and objects in the monocular field were presented to detract from the effectiveness of the convergence cue. When the real object was used in place of the reticles (Experiment IV) both types of cues were present to contribute to the perception of distance for the different convergence values. The solid line curves of Figure 4 give the distance discriminations which occurred when the accommodative and convergence cues were in agreement. In this experiment (Experiment IV) the average physical distances of the real binocular object from *S* were 6, 8, 13, and 31 feet. It can be seen from Figure 4, that, even for those *S*s who evidenced some changes in perceived distance, the magnitude of these perceived changes was considerably less than the range of physical distances required to produce the different values of convergence.

It is evident that many of the *S*s used the distance cues available in this study poorly or not at all. In general this indicates the inadequacy of convergence or of the convergence, accommodation complex as a distance cue beyond distances of one or two meters. But this is the region of the visual field in which it might be expected that experience would result in the most sensitive relation between these cue systems and perceived distance. It is evident from the initial stereoacuity measures taken on the *S*s and from the ability of the *S*s to perceive movement while continuously viewing the binocular recticle, that other binocular perceptions are much more sensitive than that resulting from absolute convergence. Thus, the sensitivity of the possible function relating absolute convergence and perceived absolute distance is not such that it can provide a theoretical basis for other binocular phenomena.

F. Summary and Conclusions

In this study, the role of convergence and accommodation in the perception of absolute distance was investigated. To measure perceived absolute distance the *S*s were provided with a visual ruler which consisted of a monocularly viewed alley containing numbered cards. A single binocular object, either stereoscopically generated or a real object, was presented to the *S*s at different convergence and accommodative distances. The *S*s judged the apparent distance of the binocular object with respect to the cards in the monocularly observed alley. The relation between convergence and perceived distance with accommodation held constant was determined for distances 0.9 meters or more from *S*. The perception of distance with convergence and accommodation varying concomitantly was determined for distances 1.8 meters or more from *S*. The results indicate that while some persons were able to use these cue systems under these conditions to some extent, in general at these distances these cues are not very effective determiners of perceived distance. Convergence, therefore cannot be an explanatory device for understanding other, more precise binocular phenomena such as stereopsis which readily occurs within the range of binocular stimulation used.

REFERENCES

1. Baird, B. A. The influence of accommodation and convergence upon the perception of depth. *Amer. J. Psychol.*, 1903, **14**, 150-200.
2. Bourbon, B. La perception visuelle de l'espace. Paris: Schleicher Freres, 1902. Cited by R. S. Woodworth, Experimental Psychology. New York: Henry Holt & Co., Inc., 1938, p. 670.
3. Carr, H. A. An Introduction to Space Perception. New York: Longmans, Green & Co., 1935, 245-259.
4. Gogel, W. C. An observer constant in the perception of stereoscopic depth. *U.S. Army Med. Res. Lab.*, Report No. 316, 1-24, Fort Knox, Ky., 1958.
5. Gogel, W. C., Hartman, B. O., & Harker, G. S. The retinal size of a familiar object as a determiner of apparent distance. *Psychol. Monog.*, 1957, **70**, 1-16.
6. Grant, V. W. Accommodation and convergence in visual space perception. *J. Exper. Psychol.*, 1942, **31**, 89-104.
7. Ittelson, W. H. The Ames Demonstrations in Perception. New Jersey: Princeton Univ. Press, 1952.
8. Kilpatrick, F. P. (Ed.). Human Behavior from the Transactional Point of View. Institute for Associated Research: Hanover, N. H., 1952.
9. Osgood, C. E. Method and Theory in Experimental Psychology. New York: Oxford Univ. Press, 1953.
10. Swensen, H. A. The relative influence of accommodation and convergence in the judgment of distance. *J. Gen. Psychol.*, 1932, **7**, 360-380.
11. Woodworth, R. S. Experimental Psychology. New York: Henry Holt & Co., Inc., 1938, 665-680.
12. Woodworth, R. S., & Schlosberg, H. Experimental Psychology. New York: Henry Holt & Co., Inc., 1954, 475-480.

Psychology Division
U.S. Army Medical Research Laboratory
Fort Knox, Kentucky

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 303-315.

CONVERGENCE AS A CUE TO THE PERCEIVED DISTANCE OF OBJECTS IN A BINOCULAR CONFIGURATION*

U.S. Army Medical Research Laboratory, Fort Knox, Kentucky

WALTER C. GOGEL[1]

A. INTRODUCTION AND PROBLEM

The convergence of the eyes when fixating a binocular object is the positioning of the eyes necessary to bring the image of the object on the fovea of each eye. The kinesthetic sensations resulting from this muscular activity are a possible cue to the distance of the object from O. This means that a possible relation exists between the convergence position of the eyes and perceived absolute distance. In a previous study a method of measuring this relation was described (4). This method consisted of providing *S* with a "ruler" formed by a monocularly observed alley containing adequate cues to distance. A binocular stimulus having a particular convergence value was also presented such that it appeared to *S* to be located in the monocular alley. The *S* judged the position of the binocular object with respect to numbered cards distributed in depth in the alley. It was found with convergence values of 4 degrees or less and accommodation held constant that most *S*s did not perceive the binocular object to be at different distances for the different convergence values. Three of the *S*s, however, made reliable though restricted distance discriminations on the basis of convergence alone. When accommodation and convergence were in agreement, the proportion of *S*s who evidenced a relation between convergence and perceived distance increased. In general, the results of this previous study indicate that the convergence to a single object is frequently an inadequate cue to perceived absolute distance.

The investigation of the role of convergence in the perception of binocular phenomenon is of continuing interest. For example, this cue system is considered to have perceptual effects in Luneburg's mathematical analysis of binocular phenomena (7, 8, 9, 10, 11). It is thought that the convergence to a single object does not result in its perceptual localization (9). With a configuration of binocular objects, however, it is postulated that convergence is a significant factor. In a recent modification of the Luneburg system it is

---

* Received in the Editorial Office on June 5, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The author wishes to thank James M. H. Gregg for his assistance in collecting and analyzing the data of this study.

proposed that the addition of a constant value of convergence to each object of a binocular configuration does not have perceptual consequences (1, 2, 3, 6). It is as though the most distant object in the configuration provided the unit of perceived distance for the configuration. One consequence of this is, for example, that the perception of the visual metric relations of a configuration of binocular objects would be modified by adding a more distant object to the configuration (3).

It is clear that direct tests of the role of convergence in the perception of distance relations need to be conducted using a configuration of binocular objects. The present study provides data relevant to this problem. The question to be investigated is whether the relation between the perceived distance and convergence to a binocular object changes as a function of the number and position of other binocular objects in the configuration.

B. Experiments I, II, and III

1. *Apparatus*

a. *The Monocular Alley.* The monocularly viewed alley which was used as a kind of ruler to determine the apparent position of the binocular object or objects is schematically illustrated in Figure 1. This alley, the floor of which was covered with a checkerboard pattern of light and dark grey rectangles, was 34.2 feet long and 4.2 feet wide and appeared to *S* to extend away from himself in distance. White vertical cardboard rectangles distributed uniformly in distance along the alley were numbered consecutively with increasing distance, with the nearest card numbered one and the farthest numbered 15. The nearest white card was 5.3 feet from *S* with 2.1 feet between cards. The moveable cart shown on the left of Figure 1 was not used in the study.

b. *The Binocular Objects.* The binocular objects (reticles) were stereoscopically generated by using a pair of reflex sights with one reflex sight before each eye of *S*. Each reflex sight contained an aperture source of light of the shape required for the particular reticle. The light from this source passed through an optical system and emerged as parallel light. This parallel light was then reflected from a reflecting-transmitting surface into an eye of *S*. The reflecting-transmitting surface in front of the right eye permitted *S* to see both the alley and the reticle with his right eye. However, a black opaque cover behind the left reflecting-transmitting surface permitted *S* to see only the reticle with his left eye. The monocular alley was never seen with the left eye. Some principles involved in the type of instrument used in this study for generating stereoscopic objects are discussed in a previous study (5).

The reticles which were used in Experiments I, II and III were designed to produce the perception of two ring-shaped luminous objects. These are schematically illustrated in Figures 1 and 2. In Experiments I and II, the upper ring appeared to be the closer object while in Experiment III this was reversed. As is indicated in Figure 2, whenever the upper or lower ring was properly fused it appeared as a single object with a short line above and below the ring. This indicated that *S* had correctly converged to the particular binocular object. The outer diameter of each ring subtended 20 minutes of visual angle with the upper ring 75 minutes of visual angle higher in the field of view than the lower ring. In order to determine the reproducibility of the experimental results, Experiments II and III were essentially

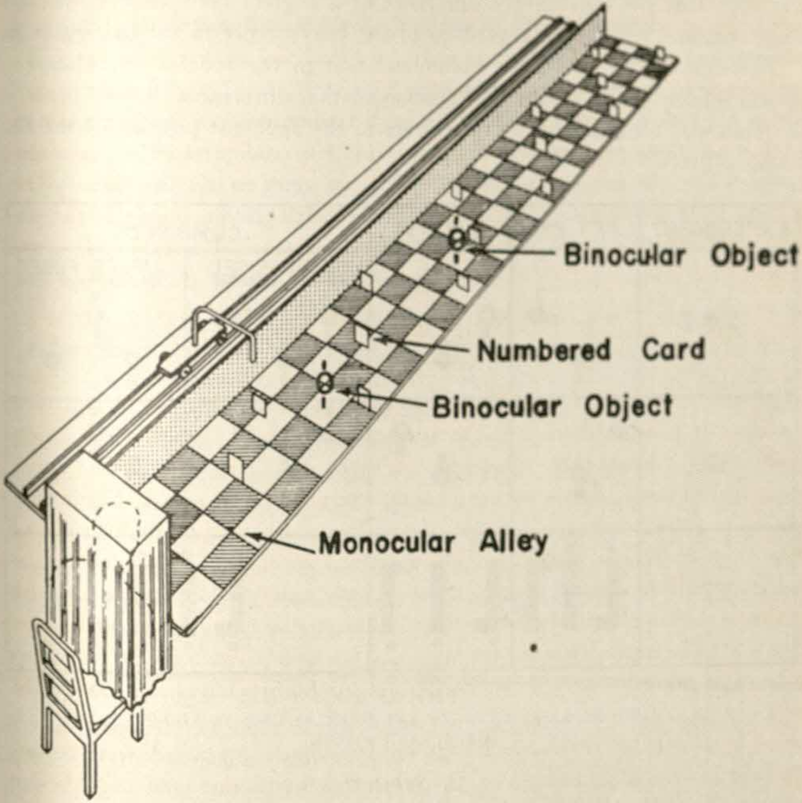


FIGURE 1
A Schematic Perspective Diagram of the Alley

repetitions of Experiment I. In all three experiments 2 degrees of convergence difference was always present between the nearer and farther binocular object. The stereoscopic instrument which produced the binocular reticles was such that the convergence difference of 2 degrees between the binocular rings remained unchanged regardless of the convergence to the configuration.

Throughout this study, the accommodation to the reticles was always to optical infinity. This means that accommodation differences did not occur in the binocular configuration to contribute to the apparent position of the binocular objects.

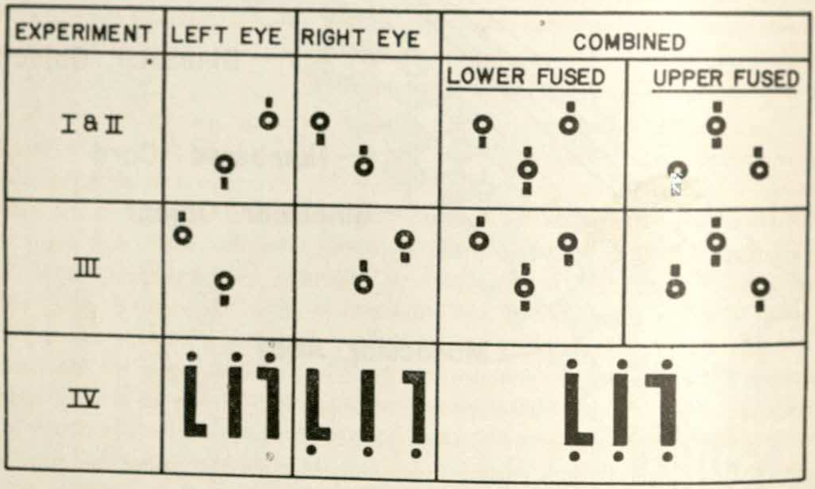


FIGURE 2
A Schematic Representation of the Reticles used to Generate the Binocular Objects.

2. *Subjects*

It was required that the *S*s be able to alternately fuse the upper and lower binocular object rapidly and at will. This was necessary in order to insure that depth was sensed between the binocular objects. Of the *S*s available only 4 were able to accomplish this and therefore only these men were used in Experiments I, II and III.

3. *Procedure*

To *S* the configuration of binocular objects appeared to be suspended above the numbered cards in the alley. The problem was to determine where in distance the binocular objects appeared to be as a function of their convergence value. For this purpose *S* judged the apparent distance position of each binocular object relative to the numbered white (monocular) cards for each of the several values of convergence. The *S*s were permitted to interpolate between cards to one-quarter card units and to extrapolate the scale of numbered cards at either end of the alley if this appeared necessary.

The values of convergence used for the most distant of the two binocular rings were 0, 48, 96, 144, 192, and 240 minutes. At each of these convergence positions, the convergence to the nearer binocular ring was 2 degrees greater, i.e., 120, 168, 216, 264, 312, and 360 minutes respectively. As used in this study, the convergence to, or the convergence value of a stereoscopic object (a binocular reticle) is the angle between the visual axes of the two eyes which results when the eyes are positioned so that the image of the object is on corresponding points on the two foveas.

The general procedure used throughout this study was as follows. The *S* was seated in the viewing position and a black velveteen cloth was adjusted over his head and shoulders. This covering plus the removal of all extraneous light sources in the room resulted in the viewing position being in complete darkness. The directional orientation of the binocular rings was adjusted so as to be unchanged in the right eye regardless of the convergence value used. The binocular objects were adjusted by *E* to the desired value, and the reticle light sources were turned on. A small black curtain which prevented *S* from seeing the monocular alley but not from seeing the binocular rings was kept lowered. Under these conditions the binocular objects appeared to *S* to be located somewhere in front of himself in the dark. The *S* was instructed as to what visual patterns would occur when the proper convergence on a binocular object had taken place. When *S* had indicated that he was able to properly fuse each of the binocular objects at will, the curtain was raised and the binocular objects appeared to be located in the monocular alley. The *S* was instructed to look from one binocular object to another fusing each in turn until he was able to estimate the depth position of each object with respect to the numbered white rectangles. Following the depth position judgments, the front curtain was lowered, the binocular reticles were turned off and the next convergence value was adjusted by *E*. This process was repeated until *S* had completed the judgments with respect to all the convergence values. Each of the six convergence values to the binocular configuration was presented three times to each *S* in random order. The order in which the *S*s judged the apparent distance of each of the objects in the configuration was systematically varied.

4. *Results*

The average reports with respect to each of the two binocular objects are shown in Figure 3 for each of the four *S*s. Since the stimuli of Experiments I and II were identical, the results from these two experiments were averaged before being combined with the results from Experiment III. In Figure 3, the curve labeled "nearer" and the curve labeled "farther" refers respectively to the nearer and farther of the two binocular objects. These curves are not substantially changed when the results from each of these three experiments are considered separately rather than averaged as they are in Figure 3. In

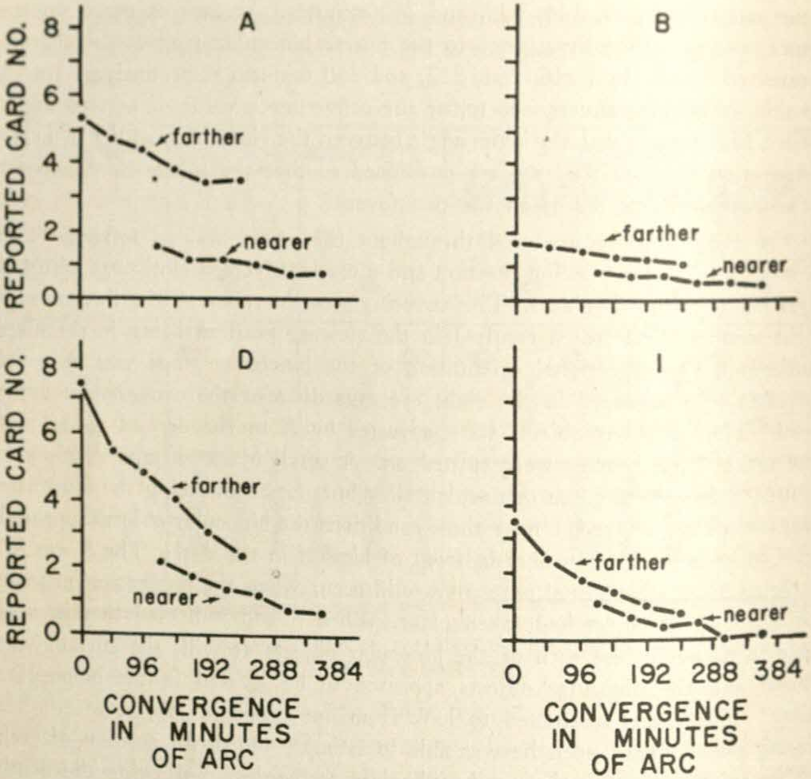


FIGURE 3
THE RELATION BETWEEN CONVERGENCE AND REPORTED DISTANCE FOR TWO BINOCULAR OBJECTS. A LARGER REPORTED CARD NUMBER INDICATES THAT THE OBJECT WAS PERCEIVED TO BE AT A GREATER DISTANCE.

this and the figures which follow, an increase in reported card number signifies an increase in perceived distance.

In all three experiments, for all four *S*s, the curves representing the reported positions of the near and far members of the pair of binocular reticles are not continuous with each other. Wherever the two curves overlap in convergence value, the reported distance position of the nearer member of the pair is less than that obtained for the farther member, when at another time is was at the same convergence. In other words, while it might be expected that the "nearer" and "farther" curves would be identical for the values of convergence at which the curves overlap, they are not (see Figure 3). This means that the perceived position of a binocular object in a configuration (in this case, a pair) of binocular objects was not independent of the configuration. From Figure 3, it would be predicted that, for a given value of convergence to a binocular object, the apparent distance position of this object would differ depending upon whether it was the nearer or more distant member of a pair of binocular objects.

It is evident from Figure 3 that *S*s A, D, and I showed some decrease in perceived absolute distance as a function of an increase in convergence. This is consistent with and in general quite similar to the results obtained previously with these same three *S*s using a single binocular object (4). This can be seen by comparing the curves for A, D, and I in Figure 3 of Reference 4 with those for A, D, and I of Figure 3 of the present study. However, it should be noted that the convergence-perceived distance relation shown by A, D, and I is not that usually obtained in situations of this type (4). The usual result is more like that of Subject B in Figure 3 in which case little or no distance discrimination as a function of convergence was evidenced.

C. Experiment IV

1. *Apparatus*

In Experiments I, II, and III only four *S*s were available who could alternately fuse the binocular objects at will. In Experiment IV, a smaller amount of binocular disparity occurred between the successive binocular objects. Therefore these objects were less difficult to fuse simultaneously or successively than in the previous experiments. As a consequence of this, more *S*s (12 men) could be used in Experiment IV. All *S*s in all four experiments, with one exception, had a visual acuity of at least 20/30 in each eye, corrected if necessary, and a stereoscopic acuity of 85 seconds of arc measured by the Diagnostic Series in the Keystone Telebinocular. The one exception had a stereoscopic acuity of 210 seconds of arc.

The reticles which were used in Experiment IV were constructed so as to produce a binocular perception of three objects at three different distances. The left binocular object appeared to be nearer than the middle binocular object which in turn appeared to be nearer than the right binocular object. As shown in Figure 2, the left and right objects were L shaped and small monocular discs appeared above and below each of the three objects. These characteristics were such that a unique configuration of stimuli was perceived when proper binocular fusion was achieved. This permitted the experimenter to determine when *S* had correctly converged to the configuration. For simplicity, each reticle object in Experiment IV will be called a stake. The convergence difference (or binocular disparity) between successive stakes was 15.5 minutes with convergence decreasing from the near left stake to the far right stake. Each stake was 7.9 minutes wide and 45 minutes high. The frontal separation between stakes was such that the overall frontal size of the reticle configuration in the right eye was 3.53 degrees. As in Experiments I, II, and III, the binocular disparity between the binocular objects in Experiment IV was constant, e.g., the binocular disparity between the stakes was always 15.5 minutes regardless of the convergence to the configuration of stakes.

In addition to the three-stake configuration of binocular objects in Experiment IV, a single binocular stake, identical to the middle stake of the three-stake configuration was also used in this experiment. The purpose of this single stake was to directly compare the relation between convergence and perceived distance obtained from a configuration of binocular objects with that obtained from a single binocular object.

2. *Procedure*

The procedure of Experiment IV was the same as that of Experiments I, II, and III. Before making the distance judgments in Experiment IV it was ascertained, when the three-stake reticle was used, that *S* saw the three stakes as being one more distant than the other in proper order. The values of convergence used for the most distant of the three stakes were 0, 48, 96, 144, 192, and 240 minutes. The corresponding values of convergence for the middle and near stake can be found by adding 15.5 and 31 minutes respectively to each of these values. With half the *S*s, the single-stake stimulus was presented before the three-stake stimulus with this order reversed for the remaining *S*s.

3. *Results*

The average results from Experiment IV are shown in Figure 4. The three curves labeled "Far," "Middle," and "Near" refer to the average judgments made with respect to the far, middle, and near object respectively in the configuration of three objects. The average results from the single binocular object are shown by the curve labeled "Single" in Figure 4. It

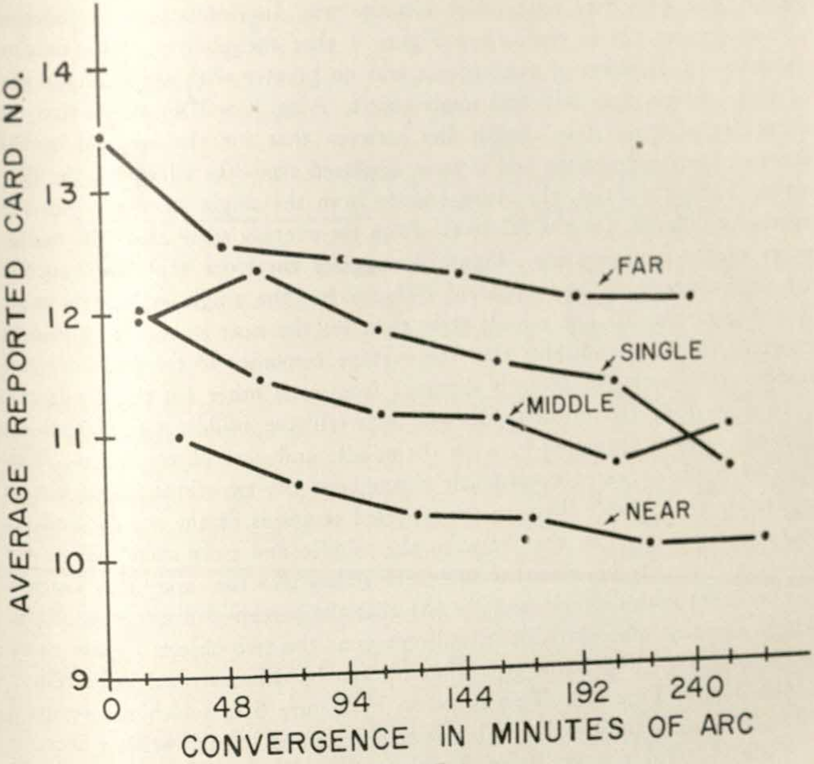


FIGURE 4
THE RELATION BETWEEN CONVERGENCE AND REPORTED DISTANCE FOR A SINGLE BINOCULAR OBJECT AND A CONFIGURATION OF THREE BINOCULAR OBJECTS. A LARGER REPORTED CARD NUMBER INDICATES THAT THE OBJECT WAS PERCEIVED TO BE AT A GREATER DISTANCE.

will be noted from Figure 4 that while there was some tendency for the average reported distance of a stake to increase with a decrease in convergence, this was slight. For example, the equivalent physical distance of the far stake ranged from infinity (0 degrees convergence) to about 90 cm (4 degrees convergence) from *S*, while the corresponding reported average change was approximately 3 feet (1.5 card units). From an examination of the results from each *S* individually, it appears that only half the *S*s showed any tendency to discriminate distance as a function of convergence. Two of these six *S*s were A and D (*S*s B and I were not used in Experiment IV).

Of particular importance in this study is a comparison of the "Single" curve with the other three curves of Figure 4. The slopes of these curves indicate the sensitivity with which distance was discriminated as a function of convergence. It is clear from Figure 4 that the ability to discriminate distance as a function of convergence was no greater with the configuration of three objects than with the single object. Also, it will be noted that the curve for a single stake usually lies between that for the far and middle stake of the configuration and is most displaced from the curve for the near stake. From the *t*-test, the average score from the single stake was not significantly different (at the .05 level) from the average score from the middle stake of the configuration. Figure 4 suggests therefore that the function relating convergence and apparent distance for the single stake was more like that for the far and middle stake than for the near stake. As would be expected the *t*-test indicated that the average responses to the near, far and middle stakes were significantly different from each other (at the .01 level).

In Figure 4, the ordinate distance between the middle and far object curves is greater than that between the middle and near object curves. This type of result appeared consistently throughout the experiment, i.e., almost invariably the distance between the reported positions of the middle and far stake was greater than that between the middle and near stake. Since the binocular disparity between the successive stakes was the same, this suggests in agreement with a previous study (5) that the perceived distance associated with a constant binocular disparity decreases as the two objects appear closer to the observer. This same type of result can be demonstrated with respect to the data of Figure 3. This is shown by Figure 5 in which the reported distance between the two binocular objects decreased linearly with a decrease in their apparent distance from *S*.

D. Discussion

It seems from these experiments that neither the sensitivity (slope) of the function relating perceived distance and convergence nor the frequency with which such a discrimination occurs with *S*s is greater when several rather than one binocular object are present. It appears for most *S*s that changing the absolute convergence to an otherwise unchanged binocular configuration had little if any effect upon either the apparent position of the configuration or the perceived distances between objects in the configuration. However, this is not the case for all *S*s. Some *S*s showed a definite but limited change in apparent distance with a change in convergence. This seemed to occur with equal sensitivity whether one or several binocular objects were involved. With these *S*s, the apparent distances between the objects in the binocular configuration seemed to be a linear function of perceived distance. Also, it was found for these same *S*s, that the curve relating perceived distance and

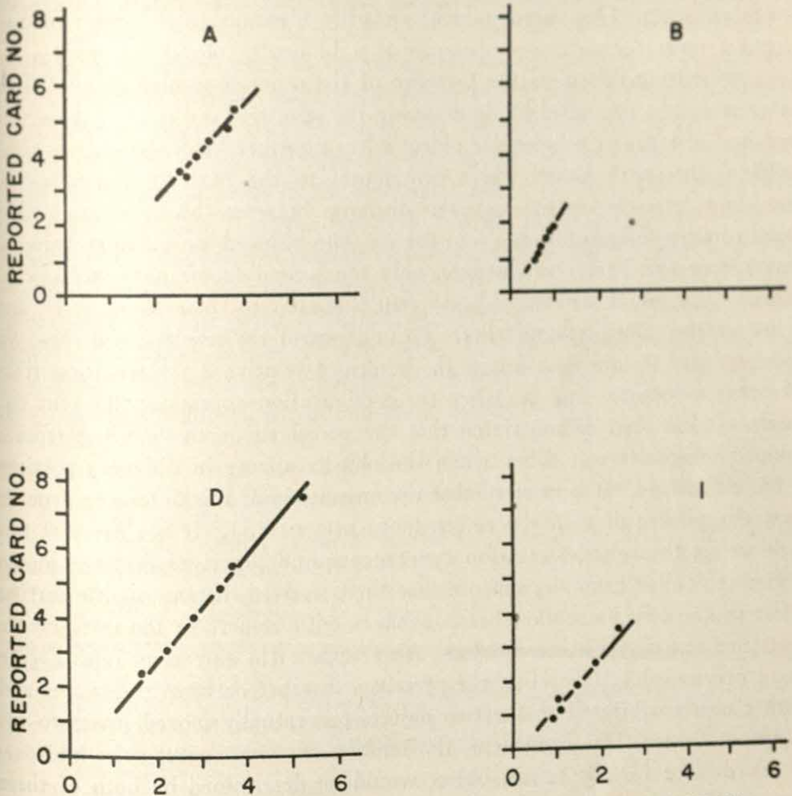


FIGURE 5
THE AVERAGE REPORTED DISTANCE BETWEEN TWO BINOCULAR OBJECTS AS A FUNCTION OF THE AVERAGE REPORTED DISTANCE OF THE FARTHER OBJECT.

convergence for a particular object was dependent upon the position of this object in the configuration of binocular objects.

The change in the relation between perceived distance and convergence for the different objects in the configuration suggests that several factors rather than a single factor were present to determine the perceived distance of each binocular object. One factor which is readily identified is the effect of convergence upon the perceived distance of a binocular object independent of any contributing effects of the presence of the other binocular objects. This factor would be measured by determining the effects of convergence upon the perceived distance of a binocular object when no other binocular objects were visible. The other factor which contributes to the perceptual localization of a binocular object is the binocular disparity between this object and other objects in the configuration. For the *S*s who showed no relation between convergence and perceived distance, only the second factor needs to be considered. But for *S*s who would give results similar to those from A, D, and I, both factors are of importance. To understand the results from these *S*s, it is necessary to understand how the perceived relative depth resulting from a binocular disparity changes when the configuration appears at different distances. It has been demonstrated that the perceived depth resulting from a binocular disparity can differ when the objects appear in different portions of the visual field. It is asserted that the magnitude of this difference depends upon the amount of frontal size constancy present (5). It is clear that this could affect the relation between convergence and perceived distance shown in Figure 3. For example, suppose that the perceived distance of the farther object of the two binocular objects behaves with respect to the convergence function like a single isolated object. Also suppose [in agreement with Figure 5 and previous data (5)] that the perceived distance between the nearer and farther object decreased as the two objects perceptually moved closer to the observer. Under these conditions, the relation between convergence and perceived distance for the nearer object would be determined by both of these factors. The usual result would be that the convergence-perceived distance relation for the nearer object would be different from that for the farther object.

From the above discussion, the amount of frontal size constancy in the measuring field (the monocular alley) can influence the convergence-perceived distance relation obtained from a configuration of binocular objects. However, it is probable that a size constancy effect of some sort will be present whether a measuring field is available or not. Hence, in most instances, it would be expected that the relation between convergence and perceived distance would differ for the different objects in the binocular configuration.

E. SUMMARY AND CONCLUSIONS

This study investigated the relation between convergence and perceived distance as a function of the number of binocular objects present. One, two, or three binocular objects were used. When several binocular objects were presented simultaneously these were at different distances from *S*. The binocular disparities between the objects in the binocular configuration remained constant as the convergence to the configuration was changed. The perceived distance of the binocular objects was measured by simultaneously presenting a monocularly viewed alley in which the binocular objects appeared to be located. The *S*s judged the apparent position of the binocular objects with respect to numbered cards in the monocular alley. It was found that neither the sensitivity (slope) of the function relating convergence and perceived distance, nor the number of persons evidencing such a relation, increased when more than one binocular object was used. For the *S*s who reported a change in perceived distance with a change in convergence, the reported distance between the objects varied linearly with the reported distance of the farther object. The obtained convergence-perceived distance function was not the same for all objects in the binocular configuration.

REFERENCES

1. BLANK, A. A. The Luneburg theory of binocular visual space. *J. Opt. Soc. Amer.*, 1953, **43**, 717-727.
2. ———. The geometry of vision. *Brit. J. of Physiol. Opt.* (3), 1957, **14**, 222-235.
3. ———. The Luneburg theory of binocular space perception. *In* S. Koch (*Ed.*), *Psychology: A Study of a Science*. Vol. 1. New York: McGraw-Hill, 1959.
4. GOGEL, W. C. Convergence as a cue to absolute distance. *J. of Psychol.*, 1961, **52**, 287-301.
5. ———. The perception of a depth interval with binocular disparity cues. *J. of Psychol.*, 1960, **50**, 257-269.
6. HARDY, L. H., RAND, G., RITTLER, M. C., BLANK, A. A., & BOEDER, P. The geometry of binocular space perception. Report of the Knapp Laboratory of Physiological Optics, Institute of Ophthalmology, Columbia Medical Center, to ONR, New York, 1953.
7. LUNEBURG, R. K. Mathematical analysis of binocular vision. Princeton, N. J. Princeton Univ. Press, 1947.
8. ———. Metric methods in binocular visual perception. *In* Studies and Essays. Courant Anniversary Volume. New York: Interscience, 1948.
9. ———. The metric of binocular visual space. *J. Opt. Soc. Amer.*, 1950, **40**, 627-642.
10. SHIPLEY, T. S. Convergence function in binocular visual space. I. A note on theory. *J. Opt. Soc. Amer.*, 1957, **47**, 795-803.
11. ———. Convergence function in binocular visual space. II. Experimental report. *J. Opt. Soc. Amer.*, 1957, **47**, 804-821.

Psychology Division
U.S. Army Medical Research Laboratory
Fort Knox, Kentucky

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 317-326.

MAGNETIC FIELDS ACCOMPANYING TRANSMISSION OF NERVE IMPULSES IN THE FROG'S SCIATIC*

Department of Psychology, University of California, Los Angeles and
The Holter Research Foundation, Inc., Helena, Montana

J. A. GENGERELLI, NORMAN J. HOLTER, AND WILFORD R. GLASSCOCK

A. THE PROBLEM

A moderately thorough survey of the neurophysiological literature during the decade 1950-60 reveals a curious lack of interest in possible magnetic accompaniments of nerve functioning. Burr (2) reviews a number of investigations on the electrical fields of biological (including neural) processes, but the center of gravity in these papers is in the electro-static rather than in the magnetic component of the field. Two articles have been found in English abstracts of Russian literature describing the effects of a constant magnetic field on nerve functioning (4, 11), and a third on the influence of electromagnetic waves of ultra-high frequency (10). A paper by Kolin *et al.* (9) reports the effects of fields of 60 and 1000 cps. Bingham (1) in surveying the period 1940-50, also reports a paucity of material devoted to the magnetic concomitants of nerve conduction.

It is regrettable that so little effort has been expended in this area, for it is not improbable that there is much to be learned here which would enrich our knowledge of neurophysiology. Events such as nerve stimulation and conduction, involving as they do the motion of electrically charged particles, must be characterized by changing magnetic fluxes. Such magnetic phenomena, if detected and studied, would in turn throw additional light on the molecular processes which generate them.

The purpose of the present report is to describe some qualitative findings that have been obtained with the sciatic nerve of the giant bullfrog. Similar experiments were reported in 1942-43 but yielded inconclusive results (5, 6).

B. APPARATUS

Two coils were fabricated for detecting the presence of magnetic fields. One of these was a toroid constituted of 12,000 turns of #46 formvar coated

---

* Received in the Editorial Office on June 5, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

wire, center tapped and wound around a supermalloy core. The core was in the form of an annulus: .5″ ID, .625″ OD and .125″ high.[1,2]

The DC resistance of the coil was 3200 ohms, and its inductance at 1000 cycles 170 henries. The coil was imbedded in epoxy resin[3] mold, to form an annulus; 2/16″ ID, 13/16″ OD and 5/16″ high.

The second coil comprised 10,000 turns of #44 wire wound around one side of a rectangular core, to form the secondary of a transformer. The coil was center tapped and insulated from the core by means of a single layer of #38 insulated wire, provided with a grounding lead. The core was made of a stack of numetal plates and formed a rectangle: inner dimensions 14/16″ × 7/16″, outer dimensions 1 12/32″ × 1″, and 4/16″ high. The DC resistance of the coil was 2600 ohms, and its inductance at 1000 cycles 291 henries. The turns of wire formed a compact spool which was treated with a heavy layer of insulating cement and hermetically sealed. The spool was on one of the long sides of the rectangular core.

The output from the coils was led to a Type 122 Tektronics differential pre-amplifier. The lead from the center tap was coupled to the ground pin of the instrument, thus protecting the signal from electrostatic artifact. The pre-amplifier's filters were set at 80 and 10,000 cycles, lower and upper bounds, respectively. Maximum gain (1000 ×) was used. The pre-amplifier output was passed through an inductance of 170 henries and a capacitance of 10 uuF, arranged in series and fed into a single beam type 208 Dumont cathode ray oscillograph. The oscillograph was operated at maximal gain. The filter provided by the inductance and capacitance served to suppress the random noise which appeared when maximal amplifications were used.

The pick-up coils, nerve preparation and the stimulation chamber were housed in a numetal box 1′ × 1′ × 1′. Signals from the coil were led from the box to the pre-amplifier over a shielded cable; power from the stimulator was led to the box over a similar shielded cable. The numetal box and the shields of the two cables were coupled to a common ground. The system proved to be very effectively shielded, and provided a clean base line on the oscillograph when both pre-amplifier and oscillograph were operating at maximal amplifications.

The toroidal coil, with amplification maximal, detected a 500 cycle current

---

[1] Code number 4T8043-S-1; Arnold Engineering Company.

[2] The authors wish to express their thanks to Coast Coil Company of Los Angeles and to Mr. C. Harris Adams for his supervision in the construction of this toroidal coil.

[3] #CRP235; Minnesota Mining and Manufacturing Company.

of about .015 uA passing through its center. The "transformer" coil gave maximal sensitivity at 1600 cycles and a current of 3 uA at that frequency was detectable with the wire passing at right angles through the plane of the rectangular core.

The stimulation chamber (Figure 1) was a lucite box 7″ × 3 10/16″ × 1 2/16″. The top of the box was fitted with a grid of stainless steel wires threaded through the upper edges of the walls and spaced one centimeter apart. At each end was a pair of wires 2 millimeters apart (S and S′) which served as stimulating electrodes. The nerve rested on this wire grid. Three of the wires nearest the stimulating electrodes were coupled together and grounded in order to minimize stimulus artifact. This ground, however, could be broken at will, as the experiment required (see Figure 1). The bottom of the chamber was covered with blotting paper soaked in Ringers, and the chamber itself rested on the bottom of the numetal box. The coil being used was held in the desired position relative to the nerve by means of a manipulator, also inside the box. After the appropriate adjustments had been made, the box was closed by its well-fitting lid (numetal).

C. Method

The sciatic nerve was severed at the exits from the vertebral column and in most instances provided a length of approximately 10 centimeters. It was used almost immediately after dissection. The preparation was placed on the grid of the chamber, with one end resting on the stimulating electrodes.

The stimulus was provided by a Grass stimulator Type S4B, monophasic adjustment, but with no isolation unit. Duration of the rectangular pulses was set at .1 millisecond and a frequency of 100 pulses per second was uniformly used. Intensity of required stimulation ranged between .9 and 2 volts. The oscillograph sweep was synchronized with the stimulator.

After each experimental session the toroidal coil was tested for electrical leakage by immersing it in Ringers and testing for continuity between one of the leads and the fluid.

1. *Experiments With the Toroidal Coil*

When the toroidal coil was used, the nerve was threaded through its center, and the coil lowered between the two wires PP′ (see Figure 1) until the nerve rested on the grid. The ground coupling was broken for the grid wires on the far side of the coil relative to the stimulating electrodes.

When the experiment required it, leads could be attached to the wires PP′ and a recording made of the nerve impulse from the region of the nerve oc-

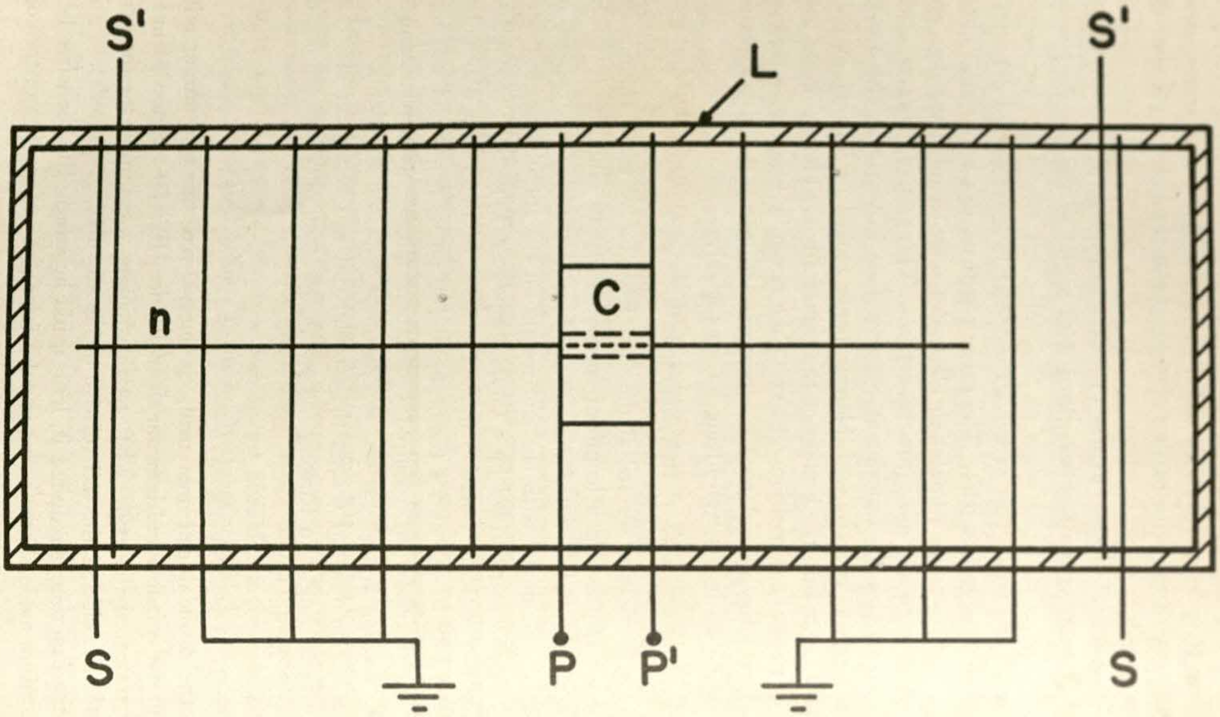


FIGURE 1
SCHEMATIC DIAGRAM OF STIMULATION CHAMBER
C, coil; PP′, pick-up electrodes; SS′, stimulation electrodes; *n*, nerve; L, walls of lucite box.

cupied by the coil. By photographing both records on the same frame, comparisons were possible. In other experiments, by pulling the nerve through the stationary coil, it could be stimulated at either end at will, thus making it possible to compare recordings obtained from the coil as the nerve impulse approached it from opposite directions.

2. *Experiments with the Transformer Coil*

Experiments with the transformer coil paralleled those done with the toroid. In this case, the side of the rectangular core opposite that bearing the windings was lowered to touch the nerve. By means of the manipulator it was possible to orient the core in such wise that the side in question rested parallel to the nerve, or by rotating it, to form any desired angle relative to the nerve's long axis. It was thus possible to determine whether the orientation of the coil was related to the signal detected. The side of the numetal core making contact with the nerve was wound with several layers of insulating plastic tape, to prevent contact of the numetal with the surface of the nerve.

Since the transformer coil was provided with an electrostatic shield, it was also possible to make observations with and without grounding of the shield.

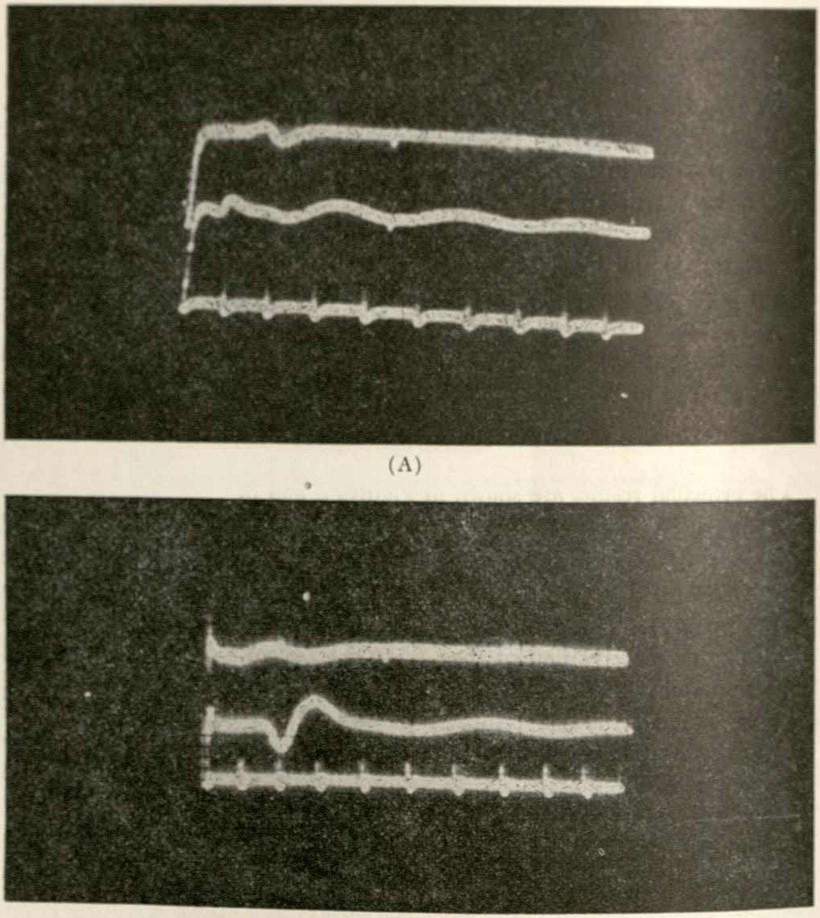
D. RESULTS

The principal results are shown in Figures 2 and 3.

In Figure 2A, the upper tracing gives the signal picked up by electrodes PP′ and shows the conventional diphasic pulse. The second tracing shows the signal transmitted by the toroidal coil. (Amplification was much smaller in the top tracing). The signal is a monophasic pip followed by mild modulations intrinsic to activation of the coil by a transient. The most interesting thing, perhaps, is the fact the signal given by the coil precedes the nerve impulse by something like .5 milliseconds. This was a constant feature of all our observations, and while a systematic study was not made, the discrepancy between the two signals tended to increase with the age and fatigue of the preparation.

In Figure 2B, the same experiment is performed with the transformer coil. The plane of the coil is at right angles to the long axis of the nerve and is centered between the electrodes PP′. Here the upper tracing represents the signal from the coil and the second, that yielded by PP′. The coil's electrostatic shield has been grounded in this instance; when the shield is not grounded, the signal is slightly stronger. Changing the orientation of the coil did not affect to any marked degree the strength of the signal, and affected its polarity not at all. The results in 2B are similar to those in 2A in that the signal from the coil precedes that from the conventional electrodes by about .5 milliseconds.

It was determined that a signal of comparable magnitude to that delivered

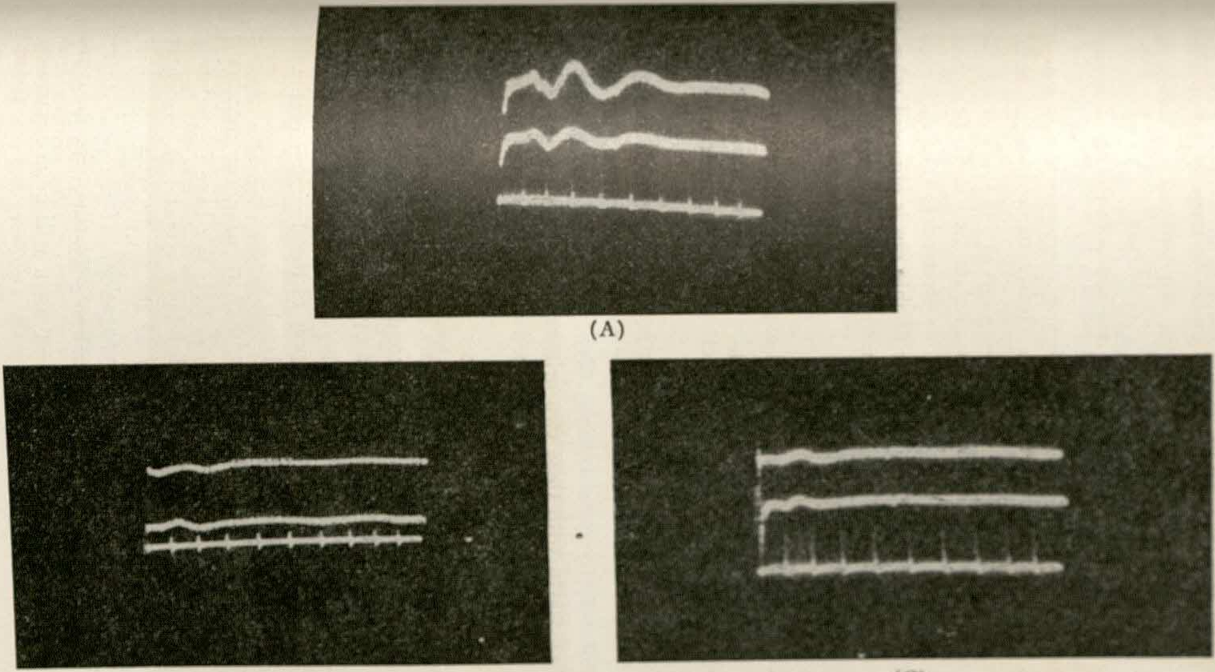


(A)

(B)

FIGURE 2

A. Comparison of tracings obtained with conventional electrodes (top line) and with toroidal coil (middle line). Coil and pickup electrodes at same distance from stimulating electrodes. Bottom line: time in this and other figures, in milliseconds. *B*. Comparison of tracings obtained with transformer coil (top line) and conventional electrodes (middle line).



(A)

(B)

(C)

FIGURE 3

A. Tracings obtained with toroidal coil, nerve impulse approaching from opposite directions in the two tracings. Top, stimulus electrode at distal end of nerve; middle, stimulus electrode at proximal end of nerve. (See text.) *B*. Tracings obtained with transformer coil. Top, as in *A*; middle, as in *A*. The plane of the core was at right angles to the long axis of nerve. (See text.) *C*. Tracings obtained with transformer coil. Top, plane of coil parallel to long axis of nerve; middle, plane of coil at right angles.

by the toroidal coil was obtained by passing through the center of the coil at 500 cycles a current of ca. .3 uA.

In Figure 3A are shown signals yielded by the toroidal coil. The top tracing was obtained by stimulating the distal end of the sciatic nerve (the end penetrating the muscles of the leg); the second, by stimulating the proximal end. In these two tracings the impulse is approaching the coil from opposite directions. Nonetheless, the pip is clearly monophasic in both instances and has the same polarity. The double pip in the top tracing relates to the fact that at the distal end of the sciatic its two large branches are separated and rest side by side on the stimulus electrodes; thus each receives the full strength of the stimulus. A similar double pip is obtained when the proximal end of the nerve is stimulated if stimulus strength is made sufficiently great. The undulations after the pip reflect the self-inductance characteristics of the toroid.

Figure 3B gives the results of an identical experiment with the transformer coil. The plane of the coil was at right angles to the long axis of the nerve and its electrostatic shield was grounded. Here we observe again no difference in the polarity of the pips.

In Figure 3C are shown results of an experiment in which the plane of the transformer coil was parallel to the long axis of the nerve (top tracing) and then at right angles (second tracing). No marked differences were found in the signals associated with these two orientations, although the signal obtained with the right angle orientation seemed generally briefer and sharper.

E. Discussion

Two results stand out in these experiments. (1) The signal delivered by the coils precedes by the order of .5 milliseconds the signal delivered by conventional electrodes when an impulse passes over a given segment of nerve. (2) The signal delivered by the coils is not perceptibly altered by varying the geometrical orientation of the coils with respect to the direction of the nerve impulse. These two phenomena must be related.

It is well known that the electrical conductivity of nerve membrane is altered at the zone occupied by the pulse (3, 7, 8). We have formulated the hypothesis that the nerve impulse is preceded by a gometrical reorientation of the membrane molecules, to permit flow of the local action currents responsible for the phenomenon of nerve propagation. If we posit that the molecular units constituting the nerve membrane are electrically charged, any motion of these, be it of translation or rotation, will be accompanied by magnetic lines of force.

It is unlikely that the magnetic lines of force emerging from the surface of the axon are mathematically symmetrical, thereby imposing equal and opposite electromotive forces upon the detectors. We should not expect completely symmetrical lattice structure in organic tissue at the molecular lever; further, the axone itself is not a perfect cylinder, nor are the axones constituting a nerve trunk perfectly collinear. If we permit ourselves to visualize the situation in the region of the fibre immediately preceding the zone of depolarization, we should imagine the segment of an imperfect cylinder with magnetic lines of force emanating everywhere on its surface very much as the quills on an excited porcupine.

These considerations would lead us to expect that (1) the signal from the coil would precede the signal from conventional electrodes, (2) the orientation of the coil with respect to the nerve would have relatively little effect on the polarity of the signal, and (3) that the signal would be monophasic.

There is, of course, the possibility that the signals detected by the system are electro-static rather than magnetic in character, and it should be admitted that, in spite of the precautions taken, this possibility remains quite real. It should be pointed out, however, that were this the case, the detected signal should have been in the form of two monophasic pulses, as reported in previous work (6). Relevant in this context, perhaps, is the fact that the present two coils were the last in a series of six or seven that was built and tried. The others yielded no results, although as electrostatic probes they should have been adequate. Electrical tests showed them to be far less sensitive to changing magnetic fluxes than the two coils herein described.

F. SUMMARY

1. Two detection coils, one with 12,000 turns of #46 wire, the other with 10,000 turns of #44 wire, were wound about highly permeable cores and utilized to detect the existence of magnetic fields around a propagated nerve impulse. The sciatic of the giant bull-frog was used.

2. Both coils delivered signals, monophasic in nature, which preceded by approximately .5 milliseconds the impulse detected by conventional electrodes located at the same region of the nerve.

3. The polarity of the signals was not affected by the orientation of the coils relative to the nerve.

4. An hypothesis is offered to explain the findings.

References

1. Bingham, Jr., W. E. Electromagnetic and electrostatic fields: a neglected area in physiological psychology. *J. Psychol.*, 1954, **37**, 225-231.
2. Burr, H. S. Design in the nervous system. *Anat. Rec.*, 1958, **131**, 405-415.
3. Eccles, J. C. The Physiology of Nerve Cells. Baltimore: The Johns Hopkins Press, 1957.
4. Erdman, G. M. Effect of a constant magnetic field on nerve. *Trud. Inst. Biofiz. Akad. Nauk SSSR.*, 1955, **1**, 189-190.
5. Gengerelli, J. A. External field of the nerve impulse. *Proc. Soc. Exp. Biol. & Med.*, 1942, **51**, 189-190.
6. ———. Electrical field of the nerve impulse. *Proc. Soc. Exp. Biol. & Med.*, 1943, **52**, 189-190.
7. Hodgkin, A. L., & Katz, B. The effect of sodium ions on the electrical activity of the giant axon of the squid. *J. Physiol.*, 1949, **108**, 37-77.
8. Hodgkin, A. L. The ionic basis of electrical activity. *Biol. Rev.*, 1951, **26**, 339-409.
9. Kolin, A., Brill, N. Q., & Broberg, P. J. Stimulation of irritable tissues by means of an alternating magnetic field. *Proc. Soc. Exp. Biol.*, 1959, **102**, 251-252.
10. Prontova, T. N. The effect of a continuous ultra-high frequency electric field on higher nervous activities in dogs under normal and pathological conditions. *Zhur. Vyssh. Neryn. Devatel*, 1956, **5**, 846-854.
11. Selivanova, E. V., *et al.* Effect of a steady magnetic field on the phenomenon of Sechenov's inhibition. *Biofiz.*, 1956, **1**, 412-415.

Department of Psychology
University of California
Los Angeles, California

The Holter Research Foundation, Inc.
Helena, Montana

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 327-331.

THE EFFECT OF ANALGESIC DOSAGES OF ACETYLSALICYLIC ACID UPON PSYCHOLOGICAL REACTION TIME*

Department of Psychology, Marymount College

MARGARET KATHLEEN CAPPONE[1]

A. INTRODUCTION

Very little attention from a psychological point of view has been devoted to the effects of acetylsalicylic acid (aspirin) upon psychological reaction time. One of the earliest studies dealing with the effects of aspirin upon reaction time was that of Macht, Issacs, and Greenberg (1918) who felt that a "full therapeutic dose" of aspirin did impair psychological reaction time. A further investigation by Davis (1936) dealing with the effect of analgesic dosages of aspirin upon a sampling of mental and motor performances, including reaction time, found that as a whole a medium (10 gr.) dosage produces practically no effect upon reaction time. In Davis' (1936) study 33 University students served for three experimental hours, once with no dose, once with a 10 gr. dosage of aspirin, and once with a control dose. In each of the three testing situations all tasks of mental and motor performance were employed. Both the aspirin and the control dose were administered in gelatin capsule form. In order to neutralize the effects of learning, the order of the doses was varied so that equal numbers of subjects began with each of the three conditions. The testing began approximately five minutes after the dose was taken and lasted approximately forty-five minutes. The reaction time test, one of Davis' (1936) tasks of mental and motor performance, consisted in the subject reacting to the click of a key. In analyzing his data Davis (1936) discarded the five most extreme reaction times in either direction and computed an average for the remaining twenty-five trials. Standard deviations were also obtained from the data. The

---

* Received in the Editorial Office on June 7, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was submitted in partial fulfillment of the requirements for the degree of Bachelor of Arts in the Department of Psychology at Marymount College, Tarrytown, New York. The author is most grateful to Doctor Mary LaVerne LaDrière, Assistant Professor of Psychology at Marymount College, Tarrytown, New York, and to Father Richard T. Zegers, S.J., Associate Professor of Psychology in the Graduate School of Psychology at Fordham University, Bronx, New York for their invaluable help in the development and successful completion of this study.

differences between performances under the three conditions were presented (for the group) in terms of rank-order correlations, and the reliabilities of the differences by the standard error of the differences.

It appears that Davis (1936) based his procedure upon the fact that since salicyl is found in the urine after only a few minutes (Hanzlik, 1926) it might be expected that any psychological effect due to its ingestion should appear within this time. However, since it takes aspirin an hour to realize its full effect (McCauley, 1960) it would stand to reason that all tasks of mental and motor performance should not have been present for each hour. Testing one task over the three hour period would have provided a more adequate estimate of the effect of aspirin in that the length of time needed for its full effect would have been taken into consideration, and consequently, the number of trials would have been increased. The statistics which Davis (1936) presented for the RT test included means, standard deviations, and rank-order correlations, on the basis of twenty-five trials gathered for each of the three experimental conditions. Utilization of such a small sample of RT responses leads one to question the reliability of the standard deviations as measures of variability for the various conditions. Further, these statistics fail to analyze thoroughly many of the possible implications in the data. Finally, Davis (1936) does not make any mention of controls for fatigue, practice prior to experimentation, suggestibility, or any criteria for the screening of subjects.

The present experiment attempts to extend the work of Davis (1936) by incorporating a more thorough experimental design and statistical procedures into the study of the effect of aspirin upon motor performance as measured by the single specific test of reaction time. Accordingly, it was hypothesized that there should not be any change in the performance of six subjects in the simple reaction time test to a white light stimulus while under the influence of any one of the four experimental conditions (5 gr., 10 gr., no dose, and placebo).

B. Apparatus and Procedure

The problem of this experiment was to determine the effect of analgesic dosages of aspirin upon reaction time. The simple Stoelting reaction time device supplied the apparatus and consisted of a chronoscope calibrated in hundredths of a second, a telegraph or response key and a white stimulus light. In extending Davis' (1936) study in design and statistical procedures the following were adopted. The criteria for selection of subjects included: (1) no use of drugs four weeks prior to actual testing; (2) no regular use of aspirin; (3) no known allergy to aspirin. Such information was ascertained by use of a questionnaire. Six female college students between eighteen and nineteen years of age were used as subjects. Each *S* served for eight experimental hours, two each with 10 gr. of aspirin, with 5 gr., placebos, and capsule. In an effort to neutralize the occurrence of any practice effect within the data and extend the reliability of the results the four experimental conditions were replicated twice and administered according to a random matrix. Since suggestibility on the part of the *E* and the *S* could play a major role in distorting the data the double-blind technique was utilized, i.e., neither the *E* nor *S* was aware of the condition of dosage since the random matrix was prepared by an individual outside of the experimental situation. The aspirin and placebos were administered in gelatin capsules in a completely darkened room. Each *S* received one-half hour of practice prior to the actual initiation of experimentation. In order to control the amount of food eaten prior to testing each *S* was required to refrain from any intake of food for eight hours before entering the testing situation; to control fatigue each *S* was required to sleep an average of eight hours per day. Accordingly the only time for conducting all experimentation was between 6:00 and 9:00 a.m. The testing began immediately after the dose was taken and lasted exactly one hour. Readings were taken every fifteen seconds, or 240 readings for the entire hour, with the foreperiod varying between two and four seconds. Means were computed for each individual *S* and these in turn were submitted to the analysis of variance technique in an effort to determine whether the differences, which might exist, could be attributable to a dosage effect.

C. Discussion of Results

The mean scores of RT were obtained for each of the six *S*s for each full hour of experimentation, and for the first, second, third, and last fifteen minutes of the hour for the eight hours of testing while under the influence of one of the four experimental conditions. These data are presented in Table 1.

Generally, the mean scores of RT presented in Table 1 were within the range of RT scores, i.e., .18 to .25 seconds, usually found in the literature (Woodworth & Schlosberg, 1956). A careful inspection of the data in Table 1 leads to the conclusion that there is no significant difference among the mean scores of the various subjects in regard to the various time periods of the experiment which could be attributed to the influence of the drug. However, because the data of *S* three showed somewhat more variabil-

TABLE 1
MEAN SCORES EXPRESSED IN HUNDREDTHS OF SECONDS FOR SIX SUBJECTS FOR THE FOLLOWING TIME SEGMENTS: 0-15, 16-30, 31-45, 46-60, AND THE FULL 60 MINUTES OF TESTING FOR EIGHT EXPERIMENTAL SESSIONS ILLUSTRATING THE APPROPRIATE CONDITION OF EXPERIMENTATION (10 GR., 5 GR., NO DOSE ((O)), PLACEBO ((P)) AND SESSION NUMBER

Condition	Session	Time Interval (Minutes) 0-15	16-30	31-45	46-60	Total 60
			Subject 1			
O	2	.20	.20	.21	.21	.21
	8	.19	.19	.20	.19	.19
P	6	.22	.21	.21	.20	.21
	7	.18	.19	.19	.20	.19
5	1	.21	.26	.24	.22	.24
	5	20	.20	.21	.27	.20
10	•3	.20	.20	.21	.21	.21
	4	.19	.20	.20	.20	.20
			Subject 2			
O	5	.22	.19	.20	.21	.20
	6	.22	.20	.20	.20	.20
P	4	.19	.20	.21	.21	.20
	8	.19	.18	.19	.22	.20
5	2	.20	.19	.20	.22	.20
	7	.18	.18	.19	.22	.19
10	1	.28	.21	.20	.23	.22
	3	.20	.20	.19	.20	.19
			Subject 3			
O	3	.19	.19	.19	.19	.19
	4	.17	.17	.19	.19	.18
P	6	.16	.17	.17	.17	.17
	8	.16	.18	.17	.18	.17
5	1	.23	.23	.25	.25	.24
	7	.16	.17	.19	.18	.17
10	2	.20	.19	.21	.20	.20
	5	.17	.18	.18	.18	.18
			Subject 4			
O	1	.22	.21	.21	.21	.21
	3	.20	.20	.20	.20	.20
P	2	.19	.21	.21	.22	.20
	5	.20	.20	.21	.20	.20
5	4	.19	.21	.19	.20	.20
	6	.20	.22	.22	.22	.21
10	7	.20	.20	.19	.20	.20
	8	.16	.17	.16	.16	.16
			Subject 5			
O	2	.20	.20	.20	.19	.20
	5	.18	.19	.19	.20	.19
P	6	.19	.19	.21	.20	.20
	8	.20	.20	.20	.20	.20
5	4	.19	.19	.20	.20	.19
	7	.20	.20	.19	.20	.20
10	1	.19	.19	.20	.20	.20
	3	.19	.19	.20	.20	.20

TABLE 1 (*continued*)

Condition	Session	0-15	16-30	31-45	46-60	Total 60
			Subject 6			
O	2	.22	.21	.22	.21	.21
	4	.18	.19	.19	.20	.19
P	3	.19	.18	.18	.20	.19
	6	.18	.17	.17	.17	.17
5	5	.18	.19	.20	.19	.19
	8	.17	.19	.19	.19	.18
10	1	.27	.18	.17	.17	.20
	7	.17	.17	.18	.17	.17

ity than that of the other five *S*s an analysis of variance was performed on the data of each subject. The main variables listed were time period and dosage level. None of the resulting F ratios attained significance, thus confirming the observational conclusion. Thus, the experimental conclusions stated by Davis (1936) were essentially validated. In view of the results of this experiment it appears warranted to conclude that within the limits imposed by the research design, dosages of aspirin do not cause any alteration in RT.

D. Summary and Conclusion

In order to determine the effect of analgesic dosages of aspirin upon the psychological reaction time of six female college students the simple reaction time response to a white light was used with readings being taken every fifteen seconds for one consecutive hour. Each subject served eight hours, two each with 10 gr. of aspirin, 5 gr., placebos, and no dosage.

The indications of the mean scores of RT and the results of the analyses of variance were that there was no alteration in RT for any of the six *S*s during any of the designated time segments nor for the total sixty minute period while under the influence of the four experimental conditions.

References

1. Davis, R. C. The effects of analgesic dosages of aspirin (acetylsalicylic acid) on some mental and motor performance. *J. Appl. Psychol.*, 1936, **20**, 481-487.
2. Hanzlik, P. J. Action and uses of salicylates and cincophen in medicine. *Medicine,* Baltimore, 1926, **5**, 197-373.
3. Macht, D. I., Issacs, A., & Greenberg, G. Action of some antipyretic analgesics on psychological reaction time. *Psychobiol.*, 1918, **1**, 327-338.
4. McCauley, J. The effects of acetylsalicylic acid on the pain threshold as determined by radiant heat. Unpublished Master's dissertation, Fordham University, Bronx, New York, 1960.
5. Woodworth, R. S., & Schlosberg, H. Experimental Psychology. Henry Holt & Co., New York, 1956, pp. 15-39.

121 E. Lotus Road
Wildwood Crest, New Jersey

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 333-334.

ASSIMILATION AND CONTRAST EFFECTS OF ANCHORING STIMULI ON JUDGMENTS: A PARTIAL REPLICATION OF THE SHERIF, TAUB, AND HOVLAND STUDY*

Barnard College, Columbia University, and University of Southern California

---

Lucy Bravo and M. S. Mayzner

A. Introduction

In a recent study Sherif, Taub, and Hovland (3) reported that the introduction of anchoring stimuli immediately adjacent to the stimulus range being judged will produce "assimilation," i.e., displacement of judgments in the direction of the anchor; while anchors distant from the stimulus range will produce "contrast," i.e., displacement of judgments away from the anchor. While contrast effects are well established phenomena (1, 2), assimilation effects have not been studied as intensively. The present report, by replicating certain portions of the Sherif, Taub, and Hovland study, provides additional evidence bearing on these effects.

B. Method

Employing the basic procedure of the Sherif, Taub, and Hovland study, judgments of weight were obtained by the method of single stimuli with and without anchoring stimuli. In the main experiment, in which the stimulus series consisted of weights of 55, 75, 93, 109, 125, and 141 gm., only two departures from the original study occurred: (1) *S*s lifted the weights directly, rather than using a lifting apparatus, and (2) only two anchor conditions were examined, 141 and 347 gm., providing the two anchoring extremes of the original study. In the supplementary experiment, in which the stimulus series consisted of weights of 97, 118, 137, 153, and 169 gm., only one departure from the original study occurred, namely, only two anchor conditions were examined, 97 and 41 gm., providing the two anchoring extremes previously employed.

C. Results and Discussion

The results of this study are presented in Tables 1 and 2, which show the mean number of judgments for the various no anchor and anchor con-

---

* Received in the Editorial Office on June 9, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

TABLE 1
MEAN DISTRIBUTION OF JUDGMENTS FOR SIX *S*S FOR NO ANCHOR (NA), 141 GM. ANCHOR (141A), AND 347 GM. ANCHOR (347A) CONDITIONS

Judgment categories	Mean number of judgments NA	 141A	 347A
1 (Lightest)	44	49	78
2	52	54	74
3	57	52	72
4	51	45	50
5	50	52	20
6 (Heaviest)	46	48	6

TABLE 2
MEAN DISTRIBUTION OF JUDGMENTS FOR THREE *S*S FOR NO ANCHOR (NA), 97 GM. ANCHOR (97A), AND 41 GM. ANCHOR (41A) CONDITIONS

Judgment categories	Mean number of judgments NA	 97A	 41A
1 (Lightest)	61	61	7
2	56	59	47
3	59	59	74
4	62	60	78
5 (Heaviest)	62	61	94

ditions. It can be seen that in both tables there is essentially no difference in the distribution of judgments for the no anchor condition versus the condition in which the anchor was either the heaviest weight in the original stimulus series (Table 1) or the lightest weight (Table 2). An application of the Kolmogorov-Smirnov test to these distributions yields *p*-values well above the .10 level. In contrast, a comparison of the no anchor condition versus the condition in which the anchor was either much heavier than the heaviest weight in the series, or much lighter than the lightest weight, shows a very significant difference in the distributions of judgments, with the Kolmogorov-Smirnov test yielding a *p*-value of less than .001. Thus, this experiment finds strong evidence for a contrast effect, but no evidence for an assimilation effect, which suggests that further study is required before assimilation, at least in the judgment of lifted weights, be accorded the same status as contrast effects.

REFERENCES

1. McGARVEY, H. R. Anchoring effects in the absolute judgment of verbal materials. *Arch. Psychol.,* 1943, No. 281.
2. ROGERS, S. The anchoring of absolute judgments. *Arch. Psychol.,* 1941, No. 261.
3. SHERIF, M., TAUB, D., & HOVLAND, C. I. Assimilation and contrast effects of anchoring stimuli on judgments. *J. Exp. Psychol.,* 1958, **55**, 150-155.

15 Sydney Avenue
Deal, New Jersey

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 335-345.

DEVIATIONS FROM MATCHING BEHAVIOR IN PROBABILITY LEARNING*[1]

Defence Research Medical Laboratories, Toronto

JOAN E. JONES[2]

A. INTRODUCTION

In some learning experiments, *S* is required to make repeated guesses about which of a number of events of fixed relative probability is about to occur. Gardner (6, 7) observed that, when more than two response categories are available, $P(\text{A})$, the frequency of calls of the most frequently-presented event (A), eventually exceeds $\pi(\text{A})$, its actual frequency of presentation. McCormack (9) confirmed Gardner's result in a replication of one of his conditions. Increasing the number of minor stimulus categories without increasing the number of response categories has no effect on $P(\text{A})$ (8). The latter result was also reported by Schipper (11) but he used a very short training period.

Schipper (11) originally predicted that in multiple-choice situations the perceptual enhancement of A will produce an inflation of $P(\text{A})$ when the events other than A are many but each of them rare. Schipper's and Gardner's results (8) indicate that this theory is untenable. Gardner (8) has related the effect to the extinction of the minor response categories which have lower probability of reinforcement in the multiple-choice condition. In all of Gardner's experiments $\pi(\text{A})$ was .60 or .70 but the effect of varying $\pi(\text{A})$ while holding the number of response categories constant can be predicted from his hypothesis. When $\pi(\text{A})$ is small $\pi(\text{B}) \ldots \pi(\text{N})$ will be larger than when $\pi(\text{A})$ is large so the probability of reinforcement of the minor response categories is higher. It follows that enhancement of $P(\text{A})$ is a function of $\pi(\text{A})$.

Three experiments are reported here. The first was designed to confirm the finding that $P(\text{A})$ is a function of the number of response categories when $\pi(\text{A})$ is constant and to test the hypothesis that $P(\text{A})$ is a function of $\pi(\text{A})$

---

* Received in the Editorial Office on June 12, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Defence Research Medical Laboratories Report No. 241-5, PCC No. D77-94-01-16, H.R. No. 205.

[2] The author appreciates the cooperation of the Department of Psychology, University of Toronto, in making available subjects and facilities for Experiment 3.

when the number of response categories is constant and greater than two. However, inflation of $P(A)$ did not occur even in the condition which was common to Gardner's and McCormack's experiments. An ancilliary feature of Experiment 1 was a retest of *S*s at a later date to assess the permissibility of re-using *S*s in experiments of this type. Experiment 2 investigates the possibility that the discrepancy between the present results and those of Gardner and of McCormack is due to different rates of presentation of events and in the interaction of rate with the number of response categories. Since the Gardner-McCormack result again failed to appear Experiment 3 was conducted to compare the performance of three groups of *S*s drawn from different populations.

B. EXPERIMENT 1

The design of Experiment 1 is summarized in Table 1.

TABLE 1
DESIGN OF EXPERIMENT 1

		Session 1 (naive)			Session 2 (experienced)		
GP	N	Number of stimulus categories	Number of response categories	$\pi(A)$	Number of stimulus categories	Number of response categories	$\pi(A)$
I	12	5	2	.60	5	2	.40
II	12	5	2	.40	5	2	.60
III	12	5	5	.60	5	5	.40
IV	12	5	5	.40	5	5	.60

Groups I and II were presented with a schedule of five colours and guessed whether or not a light of the most frequently presented colour would follow a white warning light. An A response was the verbal response "Yes" and a non-A response was the verbal response "No." Groups III and IV guessed which of the five coloured lights would occur and indicated their choice by use of the five appropriate verbal responses. Groups I and III were presented with an identical schedule of coloured lights, as were Groups II and IV. *S*s were re-tested on the same task but under a different schedule approximately two weeks after original testing.

1. *Schedules*

In constructing the schedules, the less frequent stimulus categories assumed equal values of π. Different orders in the 200 trial sequence were obtained for each *S* by various permutations of four independent series of 50 trials, each of which satisfied the relative frequencies specified and also the first-order conditional probabilities.

2. *Subjects*

Forty-eight women, wives of Air Force personnel, acted as *S*s, twelve being allocated to each of four groups.

3. *Apparatus*

S and *E* sat at opposite sides of a table with a large screen between them. The room was dark except for a lamp on *E*'s side of the screen. On the table in front of *S* was a 6″ by 8″ black box with a ¼″ diameter aperture in the centre. Behind the aperture was a milk-glass screen onto which lights of various colours could be projected. The colours were white, red, green, yellow, purple, and blue. The colour of the light at each presentation was selected manually by *E* and its duration (3 seconds) was controlled by an electronic timer as was the interval between stimuli (3 seconds).

4. *Procedure*

After giving the instructions, *E* presented and named the white warning light and each of the five colours. A practice series of 5 trials was given in which each of the colours appeared equally often. Then followed 200 training trials under the conditions scheduled. A single trial, comprising warning light, pause, coloured light, pause, lasted 12 seconds, the whole training session lasting 40 minutes (5 trials per minute). For one-third of the *S*s in each of the groups the most frequently presented colour was red, for a third green, and for a third yellow. The colour of the most frequent light was changed for each *S* in Session 2, conducted two weeks later.

5. *Results*

The results of Experiment 1, Session 1, are shown in Figure 1. The performance curves of the .60 and the .40 groups diverge during training. Analysis of variance based on blocks 8-10 shows the effect of schedule (π) to be significant at the .1 per cent level of confidence ($F = 21.81$, $df = 1/44$). Group I, with two response alternatives, exhibits consistent matching performance from block 6 to the end of training. Group II, where the most frequent stimulus event occurs less than half the time, approaches the matching level more slowly. The curves for Groups III and IV, with five response alternatives, still lie below those of the two-response alternative groups at the end of training although the difference on blocks 8-10 is not significant ($F = 1.90$, $df = 1/44$, $P > .05$). No interaction between the number of response alternatives and schedule appeared in the analysis ($F < 1$).

In Session 2 a similar relationship between the curves for the four conditions was obtained. The curves for all conditions except one (two-choice with $\pi(A) = .60$) lie above the corresponding curves for Session 1, but the

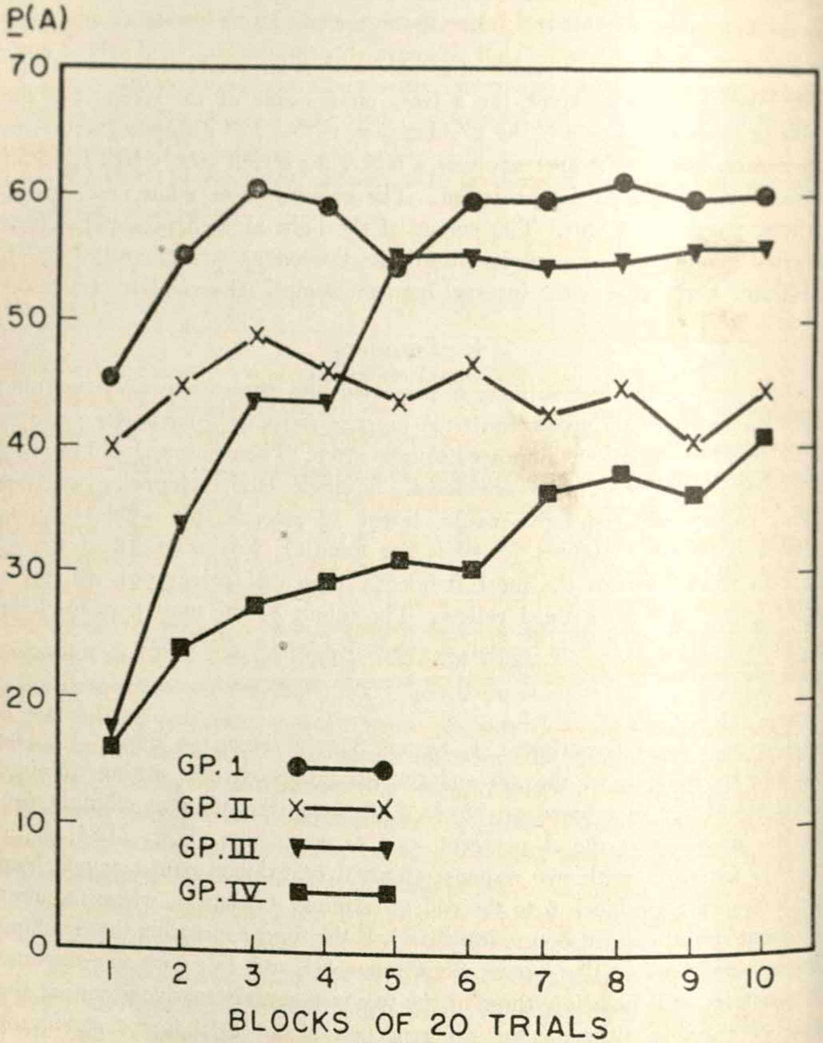


FIGURE 1
GUESSING BEHAVIOR AS A FUNCTION OF $\pi(A)$ AND THE NUMBER OF RESPONSE CATEGORIES. DATA FROM SESSION 1 SHOWING $P(A)$ ON SUCCESSIVE BLOCKS OF 20 TRIALS FOR FOUR GROUPS.

differences on blocks 8-10 are small and in no case satisfy the 5 per cent level of confidence. The difference does not necessarily indicate a more accurate probability match in experienced *S*s. Experience influences performance on the first block of trials more than on later ones, the effect apparently depending on the particular conditions experienced in Session 1. The design imposes severe limitations on this conclusion as each condition in Session 2 appears in conjunction with only one kind of experience in Session 1. Table 2 summarizes the initial $P(\text{A})$ scores for naive and for experienced *S*s.

TABLE 2
$P(\text{A})$ SCORES ON THE FIRST BLOCK OF 20 TRIALS AS A FUNCTION OF AMOUNT AND KIND OF EXPERIMENTAL EXPERIENCE. THE VALUE SHOWN IN BRACKETS FOR EXPERIENCED *S*S IS $\pi(\text{A})$ EXPERIENCED BY THEM IN SESSION 1.

	Two response alternatives		Five response alternatives	
$\pi(\text{A})$	Naive	Experienced	Naive	Experienced
.60	.455	.375 (.40)	.175	.305 (.40)
.40	.40	.555 (.60)	.165	.270 (.60)

The data from the two-response alternative groups are consistent with the results of former studies (8, 11). However, contrary to the findings of Gardner (6) and McCormack (9) no evidence of inflated $P(\text{A})$ values is found in the groups given five response alternatives. The performance curve for Group III shows no sign that the matching level would be exceeded with further training, being remarkably stable on blocks 5-10 and consistently below .60. Yet both Gardner and McCormack obtained overshooting of the presentation frequency within the same period of training under this condition. The failure to confirm the original finding makes it impossible to evaluate the hypothesis that the amount by which $P(\text{A})$ exceeds $\pi(\text{A})$ is a function of $\pi(\text{A})$.

One difference between this experiment and former ones is the rate of presentation of stimuli. Gardner set a pace of 15 trials per minute on all but the earliest trials. McCormack (10) estimates that his *S*s averaged a presentation rate of 10 trials per minute. The rate of presentation in the present experiment was only 5 trials per minute. The failure to replicate the findings of former investigators may have been due to the relatively slow rate of presentation. Experiment 2 was designed to test this hypothesis.

C. EXPERIMENT 2

Two groups (V and VI) were tested under the same conditions as Groups I and III respectively, except that trials proceeded at the rate of 10 per minute. The *S*s were 24 housewives, none of whom served in Experiment 1. Apparatus and procedure were as described for Experiment 1 except for the difference in rate of presentation, each of the four phases of a trial occupying 1.5 seconds instead of 3 seconds.

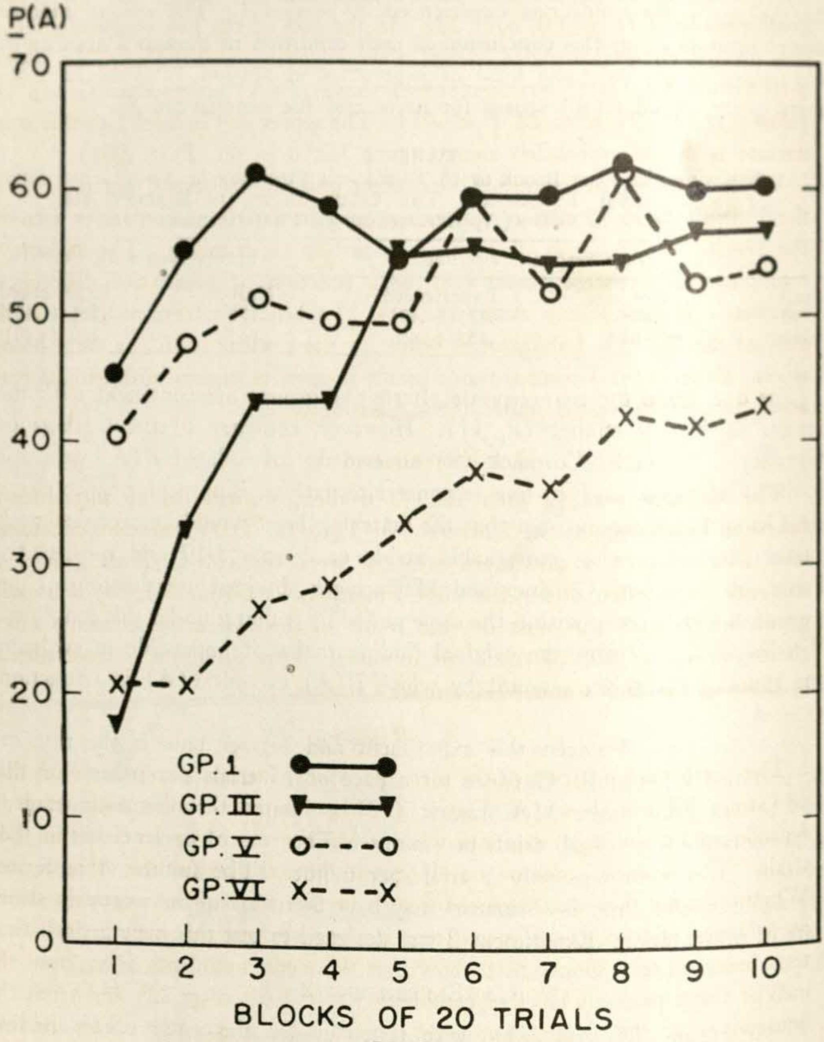


FIGURE 2

Guessing Behavior as a Function of Speed of Presentation of Stimuli—P(A) Scores on Successive Blocks of 20 Trials for Four Groups.

1. *Results*

The results of Experiment 2 are shown in Figure 2. The performance curve of the two-response alternative group run under the fast condition (V) was generally below that of the group run under the slow condition (I) but the difference on the last three blocks does not satisfy the 5 per cent confidence level ($t = 1.0$, $n = 22$). In the case of the five response alternative groups the performance curves diverge during training until on blocks 8-10 Group VI (fast) has significantly lower $P(A)$ scores than Group III (slow) ($t = 2.95$, $n = 22$, $P < .01$). The scores of Group VI at this stage deviate from the probability match ($t = 5.2$, $n = 11$, $P < .001$).

While speed of presentation does affect guessing behavior, particularly in the multiple-choice condition, it does not account for the discrepancies between the results of this study and those of earlier experiments. The failure to reproduce the expected result may be a function of population differences. Gardner's *S*s were young Army trainees, McCormack's were members of the staff of the DRML Laboratories while, in the present study, *S*s were housewives. Experiment 3 compares one of the housewife groups with groups from other populations tested under the same conditions.

D. EXPERIMENT 3

The subjects were 12 men and 12 women, students in an introductory course in Psychology at the University of Toronto. Data from two additional men who adopted a maximum strategy were discarded. Both groups of students were tested under the same conditions as Group VI, the housewife group which gave the most deviant result in the former experiments (five-choice, $\pi(A) = .60$, 10 trials per minute). This condition is most similar to those of Gardner's and McCormack's studies.

1. *Results*

The performance curves of the men and women students, together with that of Group VI are shown in Figure 3. The Gardner-McCormack result is reproduced by the men students who exceed the matching level within 200 trials. The women students barely reach the .60 level within this period. Whether or not they would exceed it with further training is an open question; no evidence of leveling-off is apparent in their performance curve. Statistical tests based on trial blocks 8-10 show that the women students differ from the men at the 5 per cent level of confidence ($t = 2.61$, $n = 22$) and from the housewives at the 1 per cent level ($t = 3.27$, $n = 22$). This finding strengthens the hypothesis that the difficulty of reproducing the Gardner-McCormack effect in the present study is a function of populations tested.

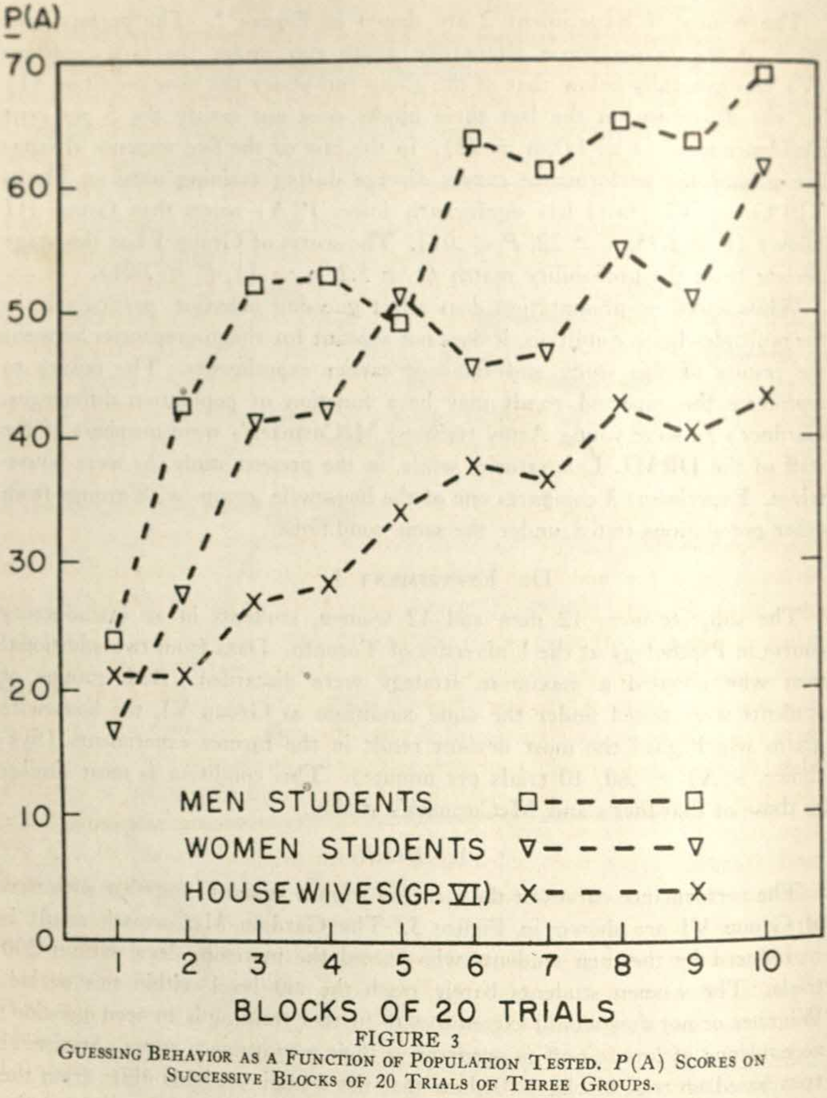


FIGURE 3

GUESSING BEHAVIOR AS A FUNCTION OF POPULATION TESTED. $P(A)$ SCORES ON SUCCESSIVE BLOCKS OF 20 TRIALS OF THREE GROUPS.

E. DISCUSSION

The question can be raised quite validly as to whether the levels of performance reached on later blocks of trials in the series of 200 presented in these experiments in fact represent or even approximate asympotes of responding. Gardner (6, 7) hesitates to accept 420 trials as sufficient to establish asymptotic performance and Edwards (2) showed that even in a two-choice situation the matching level can be exceeded if training is continued long enough. The question is particularly pertinent in the present experiments as most cases of deviation are below, rather than above, the matching level. Certainly marked differences exist between groups at this stage but whether they reflect differences in asymptotes being approached (if such exist) or rather reflect differences in the rate of approach to a common asymptote can be decided conclusively only by extending the period of testing. Using Estes' formula (3) and estimating θ from the data it was calculated that Group VI (see Figure 2) would require upward of 800 trials to reach a matching level of performance. This order of magnitude is so much greater than the number of trials required by other treatment groups in the present experiment to reach a reasonably stable level of responding that it appears unlikely that a value as high as .60 can be regarded as an appropriate asymptotic level for this condition.

It is evident that specification of the conditions under which probability matching behavior does or does not occur is not a simple matter. Though deviation from the matching level may be produced by manipulation of a given variable, both the direction and size of the deviation depend on the values assumed by other variables on that occasion.

In Experiment 1 the frequency of calls of the most frequent event by housewives is lower when five response alternatives are offered than when there are only two. In Experiment 2 this difference was much more marked when the speed of presentation was increased. In Experiment 3 the latter conditions produced a deviation in the opposite direction in a different population of *S*s (men students) i.e., with five response alternatives the most frequent event was called more frequently than it occurred. In yet another group (women students) performance was intermediate to those already mentioned.

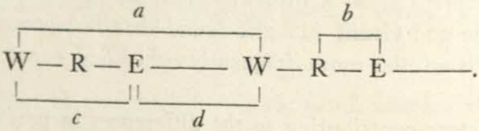
The effect of rate of presentation in Experiment 2 depends on increases in the number of response alternatives. Increasing speed of presentation depresses the frequency of calls of the most frequent event. The same effect was reported by Gardner (5) in a three-choice situation with well-differentiated stimuli. Anderson and Grant (1) also found that fast presentation produced less frequent calls of the more frequently scheduled event in a two-choice situation.

Among the factors contributing to the differences in performance between subject populations, the three curves in Figure 3 suggest that the sex of the subject is important. However, differences in academic program may be involved when men and women students are unsystematically sampled. Mc-

Cormack (9) found a small but insignificant sex difference among laboratory personnel in the same direction as that found in Experiment 3. The housewife group is differentiated from the student group in age (approximately 10 years older on the average than students), years out of school, present occupation, and very likely in amount of formal education. Fillenbaum (4) has shown that even in a population homogeneous in many respects (soldiers) very different patterns of guessing behavior may appear in sub-groups. Stevenson and Weir (12) showed that CA affects the guessing behavior of children in a three-choice probability task.

The importance of the distinction between the number of stimulus categories and the number of response categories is confirmed by the present study. The matching behavior of the two-response groups in these experiments is consistent with Schipper's (11) failure to produce deviations from the matching level when the number of stimulus events was increased and with Gardner's null result (8). In those experiments *S* predicted whether or not a given event would occur. Gardner (7) studied the effect of the number of response categories by adding "dummy" choices while presenting a constant schedule of events. This procedure produced similar but weaker effects than those produced by increasing the number of "real" choices in his 1958 experiment. Gardner related this effect to the rate of extinction of the response to rarer categories. It will be remembered that, in the populations studied by Gardner, increasing the number of minor response categories increased the frequency of calls of the most frequent event. In the housewife group in the present experiments the effect was to decrease that frequency. Thus, the extinction explanation does not fit the present case, particularly as increasing the rate of presentation further depressed the calls of the most frequent event.

It might be worthwhile to study the effect of rate of presentation. Several temporal relationships are involved but isolation of differential effects is not possible with the method used in Experiment 2. If W is the warning stimulus, R the response, and E the scheduled stimulus event, then the paradigm for two trials may be written:



The temporal intervals, *a*, *b*, *c*, and *d*, are under the control of *E*. Manipulation of *a* varies the inter-response interval and *b*, the delay of reinforcement. If these two variables affect guessing behavior a noncognitive theory of probability learning is suggested. Intervals *c* and *d*, on the other hand, relate to cognitive processes, the former being decision time and the latter "free" time between trials. Analysis along these lines might be the key to the different patterns of guessing behavior exhibited by various populations.

F. SUMMARY

Three experiments are reported in which guessing behavior in probability situations is studied as a function of $\pi(A)$, the number of response categories, the rate of presentation of events, and the population tested. The last involved differences in experimental experience, sex, age and occupation, and other confounding factors. The results suggest that the above variables interact to produce deviations from probability matching behavior, the direction and size of the deviation also depending on a combination of factors.

REFERENCES

1. ANDERSON, N. H., & GRANT, D. A. A test of a statistical learning theory model for two-choice behavior with double stimulus events. *J. Exp. Psychol.*, 1957, **54**, 305-317.
2. EDWARDS, W. Probability learning in 1000 trials. *Amer. Psychologist*, 1958, **13**, 424. (Abstract)
3. ESTES, W. K. The statistical approach to learning theory. *In* Koch, S. (*Ed.*), *Psychology: A Study of Science*, Vol. 2. New York: McGraw-Hill, 1959, 380-491.
4. FILLENBAUM, S. Matching to objective events in probability-learning—some discrepant results. *Amer. J. Psychol.*, 1960, **73**, 146-149.
5. GARDNER, R. A. Probability learning with two and three choices. *Amer. J. Psychol.*, 1957, **70**, 174-185.
6. ———. Perception of relative frequency as a function of the number of stimulus and response categories. U.S. Army Medical Research Lab. Report No. 328, 1958.
7. ———. Perception of relative frequency as a function of the number of response categories. U.S. Army Medical Research Lab. Report No. 408, 1959.
8. ———. Two-choice decision behavior with compound critical events. U.S. Army Medical Research Lab. Report No. 435, 1960.
9. McCORMACK, P. D. Spatial generalization and probability-learning in a five-choice situation. *Amer. J. Psychol.*, 1959, **72**, 135-138.
10. ———. Personal Communications, 1960.
11. SCHIPPER, L. M. Prediction of critical events in contexts of different numbers of alternative events. *J. Exp. Psychol.*, 1956, **52**, 377-380.
12. STEVENSON, H. W., & WEIR, M. W. Variables affecting children's performance in a probability learning task. *J. Exp. Psychol.*, 1959, **57**, 403-412.

Defence Research Medical Laboratories
Post Office Box 62, Postal Station "K"
Toronto 12, Ontario
Canada

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 347-361.

OPEN-MINDEDNESS AND HYPOTHESES OF INTERDEPENDENCE BETWEEN LIBERTARIAN AND EQUALITARIAN PROCESSES*

Department of Psychology, University of Arizona

SALVATORE V. ZAGONA AND WILLIAM J. MACKINNON[1]

A. INTRODUCTION

At present the world is torn by ideological conflicts which have engendered a chronic state of international tension. If permitted to continue unchecked, this tension may ignite a holocaust of unprecedented magnitude, threatening the survival of civilization itself.

It would seem that active steps taken to circumvent this consequence will involve, directly or indirectly, a re-evaluation of the social orders and ideologies involved in these differences, as well as the kinds of harmonious interrelationships possible among them.

One of the results of a careful analysis and evaluation of one's relationships with others may involve some change in his values as well as in his behavior, with reference to those relationships. Similarly, an intensive examination of the factors contributing to strife between groups, if conducted conscientiously by each group with a view to reducing tension, will probably involve some modifications in the nature of the social orders characterizing each group. It seems desirable, then, that further research be undertaken to provide some insight into the personality factors associated with the understanding and acceptance of social changes of a peaceful and democratic nature, just as research in the past has illuminated ethnocentrism, authoritarianism, and other personality variables making for susceptibility to irrational and destructive mass movements. Our intention is to consider some personality factors which may be associated with peaceful and constructive social changes —changes pointing toward conflict resolution—rather than the pathological and destructive types of forces which have been studied in the past.

Specifically we wish in the present paper to relate open-mindedness as a

---

* Received in the Editorial Office on June 23, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This paper is adapted from a doctoral dissertation completed in 1959 at the University of Arizona by Salvatore V. Zagona (1959a) under the direction of William J. MacKinnon. An abstract of the dissertation (Zagona, 1959b) and a microfilm copy are available.

personal characteristic to some particular concepts of how social change might constructively occur. The measure of open-mindedness which was selected consists of low scores on the Dogmatism Scale (Rokeach, 1956b, 1960). The particular concepts of social change used here stem from hypothesized interdependent relationships between libertarian and equalitarian processes (MacKinnon, 1958a, 1958b).

Among other things, open-mindedness can be considered as the ability to receive and process information and to respond both to one's inner self-actualizing forces and to the structural requirements of situations (Rokeach, 1960, pp. 57-58). The interdependence theory, on the other hand, hypothesizes a mutual facilitation between social orders of two different types and thus offers promise of a reduction in tension between the adherents of the two social systems.[2]

Through a study of individuals, we will attempt to relate open-mindedness to concepts of interdependence. If this relationship can be demonstrated empirically, both the sensitivity of the personality scale to the realistic information-integrating capacity of individuals and the validity of the relations hypothesized in the theory of interdependence will tend to be confirmed. In this sense we can speak of "reciprocal evaluation."[3] This evaluation comes about not only through the present study but also through prior empirical work with the personality scale (Rokeach, 1960) and the marshalling of evidence for the interdependence theory (MacKinnon, 1958a).

In a series of studies (Rokeach, 1956a, 1960; Rokeach, McGovney, and Denny, 1955), it was found that dogmatic persons solved an information-integrating problem with more difficulty than those who were found to be less dogmatic. The problem posed by these researchers was based upon the problem-solving behavior of a little bug named Joe Doodlebug, who can jump in only four directions: north, south, east and west. He cannot crawl, fly or walk, only jump; further, he cannot turn around. Once he starts in any direction he must jump four times in that direction before he can switch. A dish of food (larger than Joe) is placed three feet directly west of him. At this time he is facing north. After stopping to survey the situation, he concludes, rightly, that he must take four jumps, no more, no less, to get the food. The subject is given thirty minutes to describe the circumstances Joe must have been in to reach this conclusion.

---

[2] It is interesting to speculate that conceptualization of an interdependence between libertarian and equalitarian practices is one of the correlates or offshoots of a general scientific-humanism in its workaday form, not only because of the scales' social-emotional relevance (e.g., misanthropy) but also because of its cognitive germaneness (logic-tight compartments, etc.). The theory of open-mindedness can be found in several sources (Rokeach 1954, 1956a, 1956b, 1960).

[3] A related approach to the evaluation of ideology has been described in this journal (Rokeach, 1951).

After the subject is given these conditions, there are three new beliefs he must incorporate, one by one, which he will then integrate into the problem solution: (*a*) the facing belief (Joe does not have to face the food in order to eat it); (*b*) the direction belief (even though Joe is facing north, he can still change direction by jumping sideways or backwards); (*c*) the movement belief (when Joe stops to survey the situation he may not necessarily be a free bug).

The difficulty of this problem lies not only in the fact that the subject must overcome three sets and replace them with new ones; more important, even if he does succeed in establishing new sets he must integrate these into a new system in order to solve the problem. These discrete stages in problem-solving have been related by Rokeach to the distinction between the variables of rigidity and dogmatism: the inability to overcome sets and replace them with new ones is a function of rigidity; the inability to integrate new sets after the old sets have been overcome is a function of dogmatism. This distinction was confirmed in another study by Rokeach et al. (1955). Subjects scoring high on rigidity manifested significantly greater difficulty than subjects low on rigidity in overcoming the individual sets on the Doodlebug problem. But these two groups did not differ from each other, after the initial sets had been overcome, in their ability to integrate the new beliefs into the problem solution. Conversely, subjects scoring high and subjects scoring low in dogmatism did not differ from each other in the ease with which they overcame individual sets; however, they did differ from each other in their ability to integrate the new beliefs into the problem solution.

The authors of this ingenious experiment reason that since the Dogmatism Scale correlates about zero with intelligence, differences in cognitive functioning on the Doodlebug problem between low and high dogmatic groups cannot be attributed to differences in intelligence. It will be seen that there is a significant negative relationship between rigidity (as measured by the Gough-Sanford Rigidity Scale) and intelligence. "However," the authors argue, "that the integration measure on the Doodlebug problem fails to differentiate high from low rigid groups while at the same time differentiating high from low dogmatic groups (the latter two groups not differing from each other in intelligence) also suggests clearly that the findings bearing upon the distinction we have drawn between dogmatic and rigid thinking cannot be explained as a function of intelligence" (Rokeach et al. 1955, p. 91).

The Doodlebug problem has served as a valuable instrument in the observation of conceptual integration. In fact, various refinements and modifications of the original problem are still being made. But it cannot be denied that this procedure has two limitations: (*a*) its concepts are devoid of attitudinal-emotional factors and broader social implications; (*b*) its solution is a deductive outcome to which all can agree. By contrast, the interdependence problem may engage the individual's social values and is largely inductive. Its verification depends upon the acquisition of further evidence and the neutralization of dogmatic resistance. An analogy would be the theory of evolution before the evidence became strong, when some men evaluated it open-mindedly and others dogmatically. There is already, however, some independent evidence for the interdependence theory both at the level of small groups and of nation states (MacKinnon, 1958).

Since the scale and the theory have been partially supported by outside study, and on the assumption that the integrating ability called for in the Doodlebug problem and described in Rokeach's theory of dogmatism is related to the type of inductive integration needed to perceive the specified interdependence, a positive relation between open-mindedness and the interdependence-cognition would seem to increase the probabilities that (*a*) the Dogmatism Scale assesses integrative ability; (*b*) the theory of interdependence between libertarian and equalitarian processes indicates valid relationships between social phenomena.

In order to provide a task situation suitable for the present purpose, it is necessary that the problem be one that is social-theory oriented while at the same time demanding integration of new social processes. We would thus have a means of establishing a relationship between the conceptual integration of new social processes and other variables associated with open-mindedness. The "new social processes" were drawn from a recently advanced theory of interdependence between libertarian and equalitarian processes (MacKinnon, 1958). International in its possible application, the theory is aimed at a reduction in tensions and conflict through the mutually reinforcing nature of libertarian and equalitarian practices. (A more detailed explanation of the theory of interdependence will be presented in the next section of this paper.) It is hypothesized that a successful interdependence "response" (i.e., a conceptual integration of the two vital aspects of the theory) is related to nondogmatism or "open-mindedness" as measured by the dogmatism scale. By and large, the results of this study have tended to confirm this general hypothesis.

B. Materials, Subjects and Procedures

1. *The Dogmatism Scale and the Interdependence Solution*

A brief understanding of the meaning of the Dogmatism Scale can be obtained from consideration of its separate subscales.

The items based upon the *formal cognitive content* of dogmatism are based upon the assumption that there are certain "uniformities" in beliefs about the nature of authority, and the nature of people in general. This scale includes items which are intended to tap both general authoritarianism and intolerance. It is assumed that with an increase in dogmatism there will be an increase in the absoluteness of beliefs in positive and negative authority, and an increase in opinionation (i.e., rejection or acceptance of people depending upon whether they agree or disagree with one's own belief-disbelief system).

Items based upon the *function* of dogmatism, which seems to bear the least relationship to the cognitive organization of the social processes in our problem, tap variables having to do with personality dynamics underlying a dogmatic adherence to any ideology. These are "feelings of aloneness and isolation, anxiety over the future, self-hate, compensatory needs for self-aggrandizement, and a generally paranoid outlook on life" (Rokeach, 1956b, p. 9). (The "paranoid outlook on life" items were omitted from the final form E of the Dogmatism Scale.)

The most revealing of the dogmatism subscales for the present problem included items based upon *cognitive structure*. The significant relationship between scores on this subscale and the solution to the interdependence problem seems to be understandable in terms of the very nature of the elements within this subscale.

Cognitive structure includes items tapping (*a*) isolation within and between belief and disbelief systems; (*b*) the disbelief gradient; (*c*) relative degrees of differentiation of belief and disbelief systems, and (*d*) relation between central and peripheral parts.

The various aspects of the theory of interdependence, in the same form in which they were presented to the subjects, follow:

2. *An Exploration in Social Prediction and Evaluation*

Part I. The following pages will describe two possible changes in society. You will be asked what kind of relations might exist between these two developments. First, imagine two types of social processes which do not exist fully at the present time in the United States. Our society may or may not move in the direction of these imagined social processes.

3. *Needs–Abilities–Opportunities System*

In the Needs-Abilities-Opportunities System the following processes occur:

a. Needs and abilities which can grow if given opportunity are cultivated according to each individual's capacities or potentialities.

b. The total supply of products (goods and services) which can be used to satisfy needs or release abilities are so interchanged that each individual consumes according to his needs and produces according to his abilities.

The Needs-Abilities-Opportunities System does not imply that people are allotted objects in such a way that they become continually passive or inert, but that people are active or inactive according to their present state of needs and abilities. In the universities and other institutions of mass education, it would mean that students would receive the kind and amount of education which would fit their natural potentialities and acquired readinesses for the further development of abilities and appreciations.

4. *Beliefs–Attitudes–Elections System*

In the Beliefs-Attitudes-Elections System the following processes occur:

a. The total supply of words and other symbols which are used to transmit beliefs and attitudes about social issues are so used that every person's orientation (beliefs and attitudes on social issues) is transmitted to the same degree to all individuals oriented in like or opposite fashion.

b. The resulting beliefs and attitudes find expression in elections in which everyone holding such beliefs or attitudes participates and in which a majority vote decides the issues.

The Beliefs-Attitudes-Elections System does not refer to the number of issues or disagreements that arise but specifies a transmission of the beliefs and attitudes held by each person in proportion to the complexity of these beliefs and attitudes and also specifies voting by each such person once issues do occur. Its meaning might require that when issues spring up between groups, and communication is indirect, the newspapers and other mass media would give space or time to each belief-attitude orientation more nearly in proportion to the number of people who hold it than is now the case.

So far the descriptions of the two imagined systems have been rather abstract. To make their meaning more concrete, a list of broadly-defined goods and services will be presented. The relation of each of the two systems to these goods and services will then be discussed.

5. *Three Types of Goods and Services*

a. Bodily items: food, sex, housing, clothing, rest, exercise, transportation, preventive and curative medical care.

b. Emotional-evaluative items: contact comfort, affection, acceptance, belongingness, independence, individuality, excitement, leisure, privacy, predictability or security, variety or new experience, emotional experiencing of the values in the literature, art, philosophy, and religion of the various classes and nations.

c. Intellectual-skill items: logic, mathematics, and scientific method, content of natural and social science, self-knowledge, original thought and reason, training in reading, writing, listening, oral expression, vocational or professional preparation, interpersonal skills.

The Needs-Abilities-Opportunities System relates to each of these types of goods and services. That is, a person may have a need to consume them, an ability to produce them, or potentialities to develop such needs and abilities.

The Beliefs-Attitudes-Elections System also relates to each of these areas of goods and services. An individual may have beliefs, attitudes, and voting tendencies regarding their consumption and their production, or regarding the development of potential needs to consume them or abilities to produce them.

Next consider two types of changes of the present state of our society. These changes will be in the direction of one or the other of the imagined social processes we have described.

6. *The Needs–Abilities–Opportunities Change*

A Needs-Abilities-Opportunities Change involves a partial or incomplete change toward both processes of the Needs-Abilities-Opportunities System as related to goods and services of each of the three types.

The opposite or reverse of the Needs-Abilities-Opportunities Change would be one worded like the paragraph above except that "away from" would replace "toward."

7. *The Beliefs–Attitudes–Elections Change*

A Beliefs-Attitudes-Elections Change involves a partial or incomplete change toward both processes of the Beliefs-Attitudes-Elections System as related to goods and services of each of the three types.

The opposite or reverse of the Beliefs-Attitudes-Elections Change would be one worded like the preceding paragraph except that "away from" would replace "toward."

8. *An Exploration in Social Prediction and Evaluation*

Part II. Would you now please read all the following questions and then answer each one after due deliberation.

a. Would you explain, from your understanding of the Needs-Abilities-Opportunities System, whether it implies that all individuals would have or use the same things? Please elaborate your conceptions in two or more sentences.

b. Would you explain, from your understanding of the Beliefs-Attitudes-Elections System, whether it implies that all individuals would spend the same amount of time in transmitting their outlook to others? Please develop your ideas in two or more sentences.

9. *An Exploration in Social Prediction and Evaluation*

Part III. Any force for change may operate with or against complementary or counteractive forces. Please give for each alternative below your judgment of the likelihood of its occurrence expressed as a value from 0 to 100 inclusive. The number you write should indicate the chances (out of a hundred) that the alternative is correct. The values for the four alternatives must add to 100. However, please check to see that none of the numbers in a set of alternatives are exactly the same. That is, split any ties in one direction or the other so that each alternative in a set receives a different number. Use no fractions. Please add the numbers you insert to see that they total to 100. Scan all the questions and then record the numbers with due consideration to each question.

Number	*Alternative*
———	1. The Needs-Abilities-Opportunities Change would tend to bring about the Beliefs-Attitudes-Elections Change.
———	2. The Needs-Abilities-Opportunities Change would tend to bring about the change opposite to the Beliefs-Attitudes-Elections Change.
———	3. The Needs-Abilities-Opportunities Change would tend to prevent both the Beliefs-Attitudes-Elections Change and its opposite.
———	4. The Needs-Abilities-Opportunities Change would have no tendency either to prevent or produce the Beliefs-Attitudes-Elections Change or its opposite.
———	Total

Number	*Alternative*
———	1. The Beliefs-Attitudes-Elections Change would tend to bring about the Needs-Abilities-Opportunities Change.
———	2. The Beliefs-Attitudes-Elections Change would tend to bring about the change opposite to the Needs-Abilities-Opportunities Change.

_______ 3. The Beliefs-Attitudes-Elections Change would tend to prevent both the Needs-Abilities-Opportunities Change and its opposite.

_______ 4. The Beliefs-Attitudes-Elections Change would have no tendency either to prevent or produce the Needs-Abilities-Opportunities Change or its opposite.

_______ Total

Which of the following six statements represents your attitude? Please check one of the six blank spaces.

_______ 1. I would prefer the Needs-Abilities-Opportunities Change over the Beliefs-Attitudes-Elections Change a little.

_______ 2. I would prefer the Needs-Abilities-Opportunities Change over the Beliefs-Attitudes-Elections Change on the whole.

_______ 3. I would prefer the Needs-Abilities-Opportunities Change over the Beliefs-Attitudes-Elections Change very much.

_______ 4. I would prefer the Beliefs-Attitudes-Elections Change over the Needs-Abilities-Opportunities Change a little.

_______ 5. I would prefer the Beliefs-Attitudes-Elections Change over the Needs-Abilities-Opportunities Change on the whole.

_______ 6. I would prefer the Beliefs-Attitudes-Elections Change over the Needs-Abilities-Opportunities Change very much.

It should be noted that the terms "libertarian" and "equalitarian" were not used. The terms were omitted in order to reduce the role of sheer emotion.

10. *Subjects and Procedures*

The subjects used in this study were drawn from classes in psychology at the University of Arizona as follows:

Class	Number of subjects
The Individual in the Group	52
Personnel Psychology	24
Total	76

The subjects were instructed that they would be given a series of various kinds of tests and questionnaires; their cooperation was requested for the sake of scientific research. In order to assure complete anonymity, each subject at the outset assigned himself a "code symbol" to be used as identification on all the material he was to turn in later. The coding procedure was simple enough: each subject assigned himself a two-digit number, preceded and followed by a letter, e.g., "K 56 W." This coding device was invented with the idea that it would be rather unlikely that two or more subjects would assign themselves the same symbol; also, that it would be unlikely for the subject to forget the symbol he had assigned himself as the series of tests progressed over the weeks during the course of this study. As was anticipated, this coding technique caused not the slightest difficulty. The subjects were apparently satisfied that their responses were anonymous (as required by a number of questionnaires in the battery); no two people chose the same symbol, nor did anyone forget the symbol he had chosen.

For all phases of the testing procedure, the subjects were fully cooperative, interested in the tests, and in general well-motivated. They were assured that later they would each receive a complete listing of their scores, the descriptions of all the tests, as well as means and standard deviations for their group. In all cases the alternate-seating arrangement was followed. Finally, all instruments were scored by the writers or by a trained psychometrician.

Of the 76 subjects beginning this study, 35 were subsequently eliminated because they showed in their responses to the questions in Part II (above) that they did not understand the basic elements of the interdependence theory (i.e., they did not incorporate the new belief-disbelief system). Rokeach would probably hypothesize that inappropriate responses at this stage of the problem are at least partially a function of rigidity. But since no rigidity measure was included in our battery, this can only be left to conjecture.

Our main concern, of course, was with those subjects who were able to set aside belief-disbelief systems concerning the nature of social reality as it is presently perceived by them. The nature of the problem required, further, that they incorporate the new concepts related to the hypothetical systems (equilitarianism and libertarianism). We would then be able to observe the relationship between the dogmatism scores of these subjects and their evaluation of the relationships between these two systems.

It should be emphasized that the subjects were not previously aware of the implications of the social processes described; the term "theory of interdependence" was not referred to in their presence; in short, care was taken not to divulge the interdependent nature of the libertarian processes aside from the materials which they were asked to read.

C. Results and Discussion

1. *Explanation of Terms*

The symbol "E-L" ("equalitarian-libertarian") will refer to the questionnaire alternative stating a tendency of the Needs-Abilities-Opportunities change described in Part I (An Exploration in Social Prediction and Evaluation) to bring about the Beliefs-Attitudes-Election change. Numerical values of E-L designate the percentage of points assigned to this alternative by the subject.

Similarly, the symbol "L-E" ("libertarian-equalitarian") will refer to the tendency of the Beliefs-Attitudes-Election change to bring about the Needs-Abilities-Opportunities change; and numerical values of L-E designate the percentage of points assigned by the subject to this process.

The symbol "S" ("Sum") will refer simply to the total number of points assigned to an interdependence between libertarian and equalitarian processes (E-L + L-E).

2. *General Open-Mindedness and the Interdependence Response*

Open-mindedness scores are scores derived from the Dogmatism Scale (in reverse). That is, a high dogmatism score is a low open-mindedness score. Rank-order correlation coefficients between open-mindedness and E-L, L-E and S are reported in Table 1. The correlations reported in this paper, all of which concern one-directional hypotheses, were subjected to one-tail tests of significance.

TABLE 1
RANK ORDER CORRELATION COEFFICIENTS, OPEN-MINDEDNESS AND E-L, L-E AND S
N = 41

Variable	Rho
E-L	+.29*
L-E	+.16
S	+.34**

* $.025 < P < .05$.
** $.01 < P < .02$.

3. *Cognitive Content, Structure, and Function of Open-Mindedness and the Interdependence Solution*

Table 2 gives rank-order correlation coefficients for inverted scores on the dogmatism subscales and E-L, L-E and S. It will be noted that only the cognitive structure subscale correlated significantly with all parts of the interdependence solution.

As has been pointed out earlier in this paper, research by Rokeach and his colleagues (1955) has indicated that intelligence as traditionally measured does not contribute to the solution of the Doodlebug problem. In a later work Rokeach (1960, pp. 209-210) comments:

> It is now once more proper to ask whether the differences found could be due to differences in intelligence rather than to openness-closedness or to the experimental variable. As before, American Council on Education scores, assumed to provide a rough measure of intelligence, were available for 59 out of the 60 subjects. The results . . . do not support the hypothesis that there are differences in intelligence, either between open and closed subjects or between memory and visual subjects. Furthermore, correlation coefficients were computed between these intelligence scores and the total time required to solve the problem, the time taken to overcome the first belief, and the time taken to solve the problem after the second belief was overcome. These correlations are .08, —.16, and —.22, respectively, and are not statistically significant. Whatever differences are found in problem-solving between the closed and open groups, or between the experimental conditions, cannot be attributed to differences in intelligence. The results also suggest that intelligence, as measured, has very little to do with the way our subjects solve the Denny Doodlebug Problem, at least in the particular sample tested. These results confirm the findings reported . . . [earlier] regarding negligible correlations between intelligence and problem-solving, and a correlation of .02 between dogmatism scores and American Council on Education scores. To all of this let us add a finding by Ehrlich (1955) with Ohio State University students. He reports a —.01 correlation between dogmatism and intelligence where the latter was measured by the Ohio State Psychological Examination.

TABLE 2
RANK ORDER CORRELATION COEFFICIENTS: STRUCTURE, CONTENT AND FUNCTION OF DOGMATISM (INVERTED SCORES) AND E-L, L-E, AND S
N = 41

Dogmatism Subscale	E-L	L-E	S
Cognitive Structure	+.392[b]	+.306[a]	+.395[c]
Cognitive Content	+.226	+.207	+.184
Function	+.107	+.163	+.182

[a] $P < .05$.
[b] $P < .02$.
[c] $P < .01$.

It can be argued, then, that the open-mindedness scale taps a thinking ability not adequately measured by our standard tests. Likewise, no significant relationships were found in the present study between intelligence and the interdependence solution (Table 3). The present study further confirms the hypothesis that the Dogmatism Scale measures a mental ability not measured by conventional tests.

TABLE 3
RANK ORDER CORRELATION COEFFICIENTS, INTELLECTUAL ABILITIES SCORES AND E-L, L-E AND S
N = 41

Variable	E-L	L-E	S
Primary Mental Abilities,[a] Aptitude (2V + R)	—.14	—.11	—.12
Critical Thinking,[b] Total	—.24	—.12	—.22
Creative Thinking,[c] Total	—.18	—.07	—.14

[a] Thurstone, L. L., & Thurstone, G., 1949.
[b] Watson, G., & Glaser, E. M., 1952.
[c] Christensen, P. R., & Guilford, J. P. (Preliminary edition of manuals for administration, scoring, and interpretation of tests of creative-thinking abilities, Sheridan Supply Company, Beverly Hills, California, copyright, 1958.)

The consistent set of significant relationships found between structural open-mindedness and the L-E, E-L responses seems to suggest that the major strength of the Dogmatism Scale (like the major value of Rorschach) may lie in its structural features. Also, the greater perception of the E-L relationship may have resulted from the questionnaire's detailing of the nature of human needs. This amount of detail, which is not balanced by a corresponding amount in connection with the libertarian model, may have set off trains of thought pertaining to the mechanics of an equalitarian-libertarian relationship.

D. SUMMARY AND CONCLUSIONS

This study was primarily concerned with linking certain emotional-cognitive variables with the act of conceiving an interdependence between social processes as described in a developing social theory. Underlying the research was a feeling that the evermounting international tensions of our time require that we conduct research into (*a*) politico-economic relationships which, if properly utilized, may serve to diminish these tensions, and (*b*) human personalities capable of perceiving these pathways to international harmony.

One of our key variables was drawn from a developing theory of interdependence between libertarian and equalitarian processes, a theory which suggests a means for reducing international tensions through emphasizing the most desirable features of political-economic systems generally regarded as being incompatible. An assumption underlying the theory is that the most desirable feature in each of these systems is not only compatible with that of the other, but also interdependent (i.e., mutually strengthening).

For research purposes the theory was restructured in terms of a task situation. The subjects were briefed on the nature of the equalitarian system, then on the nature of the libertarian system. Among other questions, they were asked to indicate the nature of the relationships *between* the two systems. This task was referred to as the "interdependence problem," though this phrasing was not used in the questionnaire itself. Responses to the problem consisted in the number of points the subjects assigned to the questionnaire alternative stating that an equalitarian change would tend to bring about a libertarian change (E-L), the number of points they assigned to the alternative stating that a libertarian change would tend to bring about an equalitarian change (L-E), and a sum of these points.

The scores thus obtained were related to scores on the Dogmatism Scale and on three tests of intellectual ability. Of these the Dogmatism Scale was considered to be most important, because it taps not only some attitudinal-emotional variables (e.g., misanthropy, ethnocentrism, anxiety), but also because it is a measure of closed cognitive structure. In accordance with the basic purpose of our study, we reversed the dogmatism scores, and labeled the variable "open-mindedness."

In setting up our problem, we followed a procedure somewhat similar to that of Rokeach et al. (1955) who used the Doodlebug problem, requiring a "breaking away" from old sets (beliefs) and incorporating new ones. The *integration* of these new beliefs were prerequisite to the solution of the problem.

However, the above problem differed from that posed in the present study in at least two major respects:

1. In the Doodlebug problem, a definite solution was known to exist, whereas in the present problem, the interdependence between equalitarian and libertarian processes is only hypothesized.
2. The Doodlebug problem required the subject to reject old beliefs and assume new beliefs which were probably not linked to his attitudinal-emotional frame. The present subjects were required to integrate concepts of a broad social and economic nature.

We have seen that Dogmatism Scale scores have been related to the appropriate integration of elements in a problem requiring a cognitive organization of new belief systems. This provides us with some evidence that the type of open-mindedness measured by this scale is instrumental in conceptual integration. Similarly, some evidence has been presented which tends to support the hypotheses underlying the theory of interdependence. The present study, by using the Dogmatism Scale and the interdependence "problem" as major variables, has placed each in the position of assessing the other. As in other studies, this investigation showed that the Dogmatism Scale measures an integrative strength not related to standard measures of intellectual ability.

For the entire scale, the rank order correlation between dogmatism scores and total interdependence solution was —.34 (N = 41), significant at between .01 and .025 levels. While L-E did not reach significance (—.16), E-L (—.29) was significant at better than the .05 level. (For the purposes of this study, one-tail tests of significance were used, and the .05 level was chosen as a criterion of significance.) The *structure* subscale, which would be expected to correlate highest if the assumptions of the scale and the assumptions underlying the theory of interdependence are valid, provided the most significant relationships. Here correlations for E-L (.39), L-E (.31) and S (.39) were significant at .01, .025, and .01, respectively.

Our study has thus provided evidence tending to confirm or "positively evaluate" both open-mindedness (non-dogmatism) as a measure of realistic integration and a theory of interdependence between libertarian and equalitarian processes as an interpretation of possible avenues of social change.

References

1. MacKinnon, W. J. A theory of interdependence between libertarian and equalitarian processes and its application to international conflict. Paper read at Amer. Psychol. Assn., Washington, D.C., August, 1958. (a)
2. ———. Proposal for a project on a theory of interdependence between libertarian and equalitarian processes and its application to international conflict resolution. *Conflict Resolution*, 1958, **2**, 355-356. (b)
3. Rokeach, M. Toward the scientific evaluation of social attitudes and ideologies. *J. Psychol.*, 1951, **9**, 97-104.
4. ———. The nature and meaning of dogmatism. *Psychol. Rev.*, 1954, **61**, 194-204.
5. ———. On the unity of thought and belief. *J. Pers.*, 1956, **25**, 224-250. (a)
6. ———. Political and religious dogmatism: an alternative to the authoritarian personality. *Psychol. Monogr.*, 1956, **70**, No. 18 (Whole No. 425). (b)
7. ———. The Open and Closed Mind: Investigations Into the Nature of Belief Systems and Personality Systems. New York: Basic Books, 1960.
8. Rokeach, M., McGovney, W. C., & Denny, M. R. A distinction between dogmatic and rigid thinking. *J. Abnorm. Soc. Psychol.*, 1955, **51**, 87-93.
9. Thurstone, L. L., & Thurstone, T. G. SRA Primary Mental Abilities, Intermediate Ages 11-17, Form AA: Manual. Chicago: Science Research Associates, 1949.
10. Watson, G., & Glaser, E. M. Watson-Glaser Critical Thinking Appraisal: Manual. New York: World Book, 1952.
11. Zagona, S. V. Dogmatism and a theory of interdependence between libertarian and equalitarian prosesses: A study in reciprocal evaluation. Unpublished doctoral dissertation, University of Arizona, 1959. (a)
12. ———. Dogmatism and a theory of interdependence between libertarian and equalitarian processes: a study in reciprocal evaluation. Dissertation Abstr., 1959, No. 6 (Abstract). (b)

Department of Psychology
University of Arizona
Tucson, Arizona

Published as a separate and in *The Journal of Psychology*, 1961, 52, 363-367.

HEAD BALANCE AND SITTING POSTURE II: THE ROLE OF THE STERNOMASTOID MUSCLE*[1]

Institute for Psychological Research, Tufts University

FRANK PIERCE JONES, JOHN A. HANSON, AND FLORENCE E. GRAY

A. PROBLEM

In a previous study (3), X-ray photography was used to record differences between two erect sitting postures, an "habitual" and an "experimental," which are sharply distinguished kinesthetically. The experimental posture is brought about by changing the reflex balance of the subject's head while he moves up against gravity. It has been shown that the pattern of such a movement (as recorded by multiple-image photography) is characteristically altered from the habitual pattern (2). In the experimental posture, subjects regularly report a decrease in the feelings of weight and effort associated with the habitual posture. The X-ray study showed that the two postures differed significantly in the angle of the head to the neck, the relative height of sella turcica (and by inference the center of gravity of the head), and the distance between the first two vertebrae. It was conjectured that a dynamic balance between the weight of the head and the tonus of extensor muscles in the neck facilitates antigravity responses; that this balance is disturbed in the habitual posture; and that the operations performed in inducing the experimental posture tend to restore the normal balance.

If the hypothesis is correct, it is reasonable to expect a different behavior of muscles in the two erect sitting postures. To test the hypothesis we used surface electromyography to explore some of the muscles of the neck and trunk. In preliminary studies we found only one muscle that consistently distinguished the two postures. This was the sternomastoid (the prominent, diagonal muscle at the side of the neck). No other muscle provided so reliable an index. The data presented here were obtained later as part of a larger study of neck muscles which has been described elsewhere (4).

---

* Received in the Editorial Office on July 5, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was supported by a grant from the U.S. Public Health Service, Division of Research Grants (RG-4836).

B. SUBJECTS

The subjects were seven male students between the ages of 16 and 21. They were employed for the summer by Tufts University, six of them in maintenance work. None had any previous knowledge of the experiment.

C. PROCEDURE

The subject was seated on a flat stool. Surface electrodes were attached by Duco cement over the right sternomastoid halfway between the insertion and the sternal head; to the right upper trapezius halfway between the occiput and the seventh cervical vertebra; and to the right erector spinae two inches below the lowest point of the trapezius. A ground electrode was placed over the right scapula. Muscle potentials were amplified by an Offner Electroencephalograph. Raw muscle activity and integrator scores were recorded by an inkwriter at a paper speed of 2.5 cm. per second.

Posture was recorded as an immediate response to a specific verbal stimulus. The subject was instructed to sit in his "best" posture, in his "most comfortable" posture, and at his "greatest sitting height." He was then asked to lean forward with his lower arms resting on his knees and to move back on signal to his best sitting posture. A continuous electromyographic record was obtained for all procedures, and a photograph taken of each of the three postures.

After the habitual postures had been recorded, the balance of the subject's head was changed and he was guided into the experimental posture. Movement and posture were recorded as before.

For six of the seven subjects the series of procedures was repeated. For statistical purposes data obtained from the seventh subject were used twice to give an N of fourteen.

D. RESULTS

As measured on the photographs, mean sitting height at the seventh cervical vertebra increased from most comfortable to best and from best to greatest sitting height. There was a further slight increase when the subject was guided into the experimental posture. The seventh cervical was used as a reference point because it provides a more reliable index of height than does any point on the head, whose angle of rotation may change markedly with changes in posture (4).

Potential in the sternonmastoid increased progressively from most comfortable to best and from best to greatest sitting height. In the experimental posture, it remained on the average a little below the level of most comfortable.

The difference between the experimental and the habitual postures was not the result of changes made after the posture was assumed but was inherent in the movement from which it arose. The sample electromyograms in Figure 1 show the paterns of sternomastoid activity in one subject for each of the four postures and for the movement as well. It can be clearly seen that when he moved up in his habitual way the sternomastoid was strongly activated. In the guided movement this activity did not appear.

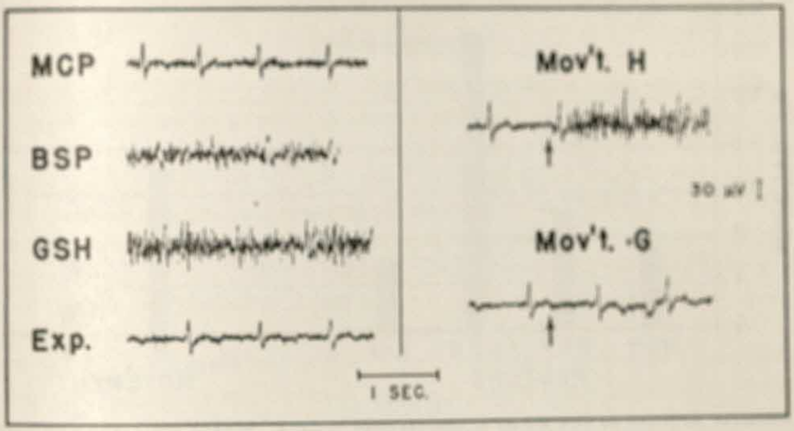


FIGURE 1
Postural Activity in the Right Sternomastoid Muscle. Sample Electromyograms From One Subject. MCP Most Comfortable (Sitting) Posture; BSP Best Sitting Posture; GSH Greatest Sitting Height; Exp Experimental Posture. H Habitual; G Guided

Similar differences between experimental and habitual postures and between guided and habitual movements were found in the means, which are shown graphically in Figure 2. All differences (experimental vs. most comfortable, experimental vs. best, experimental vs. greatest sitting height; habitual vs. guided) are significant at the .05 level or better.

E. Discussion

In the three habitual postures, the sternomastoid shows a consistent pattern of response. Its activity is lowest for most comfortable and increases rapidly as the subject increases his sitting height. (It behaves the same when the subject is standing (4)). This increase in potential does not appear in the experimental posture, which may be tentatively defined as "a condition of increased height without increased activity in the sternomastoid muscle." Subjectively, there are two hallmarks which distinguish the experimental posture: (1) to move down consciously into a slump the subject has to make an appreciable effort, as if it were necessary to overcome a positive upward force; (2) to move upward, e.g., from sitting to standing, he needs to make less than his customary effort. Both of these conditions can be observed by the experimenter if he attempts to move the subject passively. The tonus in extensor muscle seems to have increased, so that anti-gravity responses are facilitated.

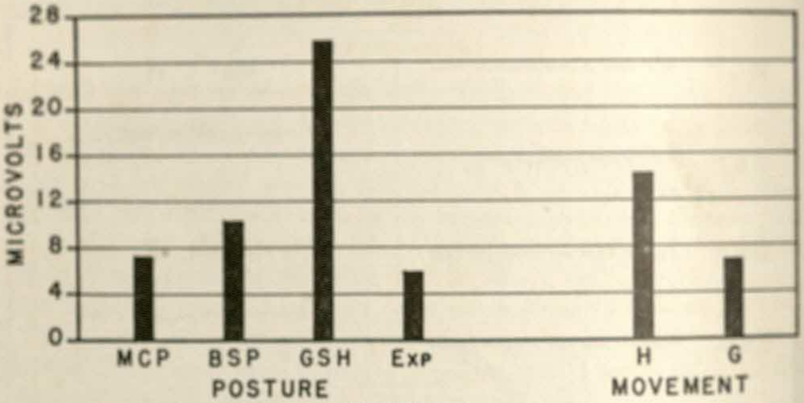


FIGURE 2
Postural Activity in the Right Sternomastoid Muscle. Means for Seven Subjects

According to Gray (1) the two sternomastoids "acting together flex the vertebral column, bringing the head forward." It seems a reasonable hypothesis that activation of these muscles is part of a flexor response which inhibits extensor tonus during the upward movement, so that the more upright posture must be maintained by voluntary effort against an active inhibitory force. The experimental procedures by eliminating this inhibition permit normal facilitation of antigravity reflexes. We submit that this hypothesis will account for the subjective phenomena and is consistent with the data previously obtained by X-ray and multiple-image photography.

F. Summary

Surface electromyography was used to study the role of the sternomastoid muscle in two erect sitting postures ("habitual" and "experimental"), which are sharply distinguished kinesthetically. In the experimental posture, which is brought about by changing the dynamic balance of the head while the posture is being assumed, the subjective feelings of weight and effort associated with the habitual posture are markedly reduced. Increased activity in the sternomastoid, which seems to be characteristic of the habitual posture, was not found in the experimental. It is hypothesized that the experimental procedures facilitate antigravity reflexes by eliminating a flexor response of the head and neck which inhibits extensor tonus.

REFERENCES

1. Gray, H. Anatomy of the Human Body. 25th ed., Philadelphia: Lea & Febiger, 1948.
2. Jones, F. P., Gray, F. E., Hanson, J. A., & O'Connell, D. N. An experimental study of the effect of head balance on patterns of posture and movement in man. *J. of Psychol.*, 1959, **47**, 247-258.
3. Jones, F. P., & Gilley, P. F. M. Head-balance and sitting posture: an X-ray analysis. *J. of Psychol.*, 1960, **49**, 289-293.
4. Jones, F. P., Gray, F. E., Hanson, J. A., & Shoop, J. N. Neckmuscle tension and the postural image. *Ergonomics*, 1961, **4**, 133-142.

Institute for Psychological Research
Tufts University
Medford 55, Massachusetts

Published as a separate and in *The Journal of Psychology*, 1961, 52, 369-378.

CONCEPT-FORMING BEHAVIOR IN SCHIZOPHRENIC AND NON-SCHIZOPHRENIC SUBJECTS*[1]

Behavioral Research Laboratory, Palo Alto VA Hospital, and Stanford University

PAUL MCREYNOLDS AND BEVERLY COLLINS

A. INTRODUCTION AND PROBLEM

Concept-forming behavior is pervasive throughout all mental life. The process whereby one assimilates his perceptions of his world can be described as the classifying or categorizing of his percepts into particular groupings according to particular conceptual schemata (1, 6, 7, 8). Such groupings—along with the schemata underlying them—are what is ordinarily meant by the term *concept*. Concept-forming behavior is thus not simply a function which the mind performs on occasion: it is an almost constant activity, the presence of which can be assumed whenever an *S* is perceiving or thinking.

Because of the ubiquity of concept-forming behavior[2] it is natural that it would have been utilized extensively in the study of schizophrenia. What is usually referred to as *thinking disturbance* or *thinking disorder* is *par excellence* an example of aberration in concept-forming behavior. The literature is too extensive to survey here but recent reviews are available (7, 10). Typically, the study of concept formation in schizophrenia has involved the presentation to the patient of stimulus objects which can be grouped to represent certain given concepts. The attempt is to evaluate the patient's *capacity* to form accurate concepts. More bluntly, the experimenter knows the correct answers and he wants to see if the patient can produce them. *E* may, of course, also be interested in how *S* goes about solving the problem and in any idiosyncrasies which *S* may show, but such observations, however useful they may be clinically, are difficult to make sufficiently objective to use as comparative data. Further, many conventional test methods are extremely limited in the range of concept-forming behaviors which they sample. These comments are particularly apt for conventional clinical concept forma-

---

* Received in the Editorial Office on July 5, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] For their generous assistance in this study the authors are indebted to Robert Weiss, Hugh Kohn, Mary Acker, Itsuye Sakai, Jack Daily, George Krieger, Richard Worthington, Robert Meiers, Richard Hamister, and Leonard Ullmann.

[2] We are using the term "concept-forming behavior" in preference to the more conventional "concept formation"—though we will use this term, too—as a way of taking account of the continuing nature of the process of categorizing percepts. "Concept formation" tends to imply a discrete, separable kind of mental capacity.

tion tests (e.g., the Vigotsky blocks). They are much less applicable to the ingenious procedures developed by some investigators (2, 3) to study particular aspects of conceptual behavior in schizophrenics.

The present study was designed to compare concept-forming behavior in schizophrenics and non-schizophrenics in such a manner as to reflect, as much as possible, typical levels of, rather than capacities for adequate concept formation. It seems possible that certain patients who customarily utilize loose standards is assigning percepts to categories may nevertheless, when faced with a given task having a correct answer, be able to mobilize their resources and solve the task, thus making characteristic patterns of conceptual behavior.

B. METHOD

1. *Subjects*

Two samples of schizophrenics were studied—those having clinically manifest thinking disturbance, and those not evincing such symptomatology. Previous studies have not made this distinction, but it is important to do so since it is a common observation among clinicians that not all schizophrenics show this kind of picture. Thus, some schizophrenics appear to be quite intact in this regard, and to have the primary manifestations of their diagnosis in other symptoms, such as delusions, hallucinations, and paranoid projections. The question arises: from the standpoint of an objective measure of level of conceptual behavior, to what extent is difficulty in this area characteristic of all schizophrenics, and to what extent is it limited to patients with clinically manifest thinking disorder?

In order to choose the schizophrenic samples a detailed definition of thinking disorder—composed largely of pertinent quotations from three standard psychiatric texts—was prepared and distributed to appropriate staff psychiatrists and psychologists at the Palo Alto Veterans Administration Hospital. These judges were asked to nominate patients who in their judgement would be among (1) the fifth of all schizophrenics in the hospital having the *most* clinical disturbance in the thinking process; and (2) the fifth of all schizophrenics in the hospital having the *least* clinical disturbance in the thinking process. It was not possible to obtain systematic estimates of the reliability of these judgments, since usually there was only one clinician who was intimately familiar with each case. However, each judge typically discussed his nominations with other staff as appropriate. A number of patients rated as most disturbed in the thinking process proved untestable due to resistiveness, bizarre behavior, and the like.

As a control group it was decided to use other hospitalized psychiatric, but non-schizophrenic patients, rather than normals, in order to control—

at least to some extent—for such variance as might be due simply to being hospitalized. Further, the important question in research of this kind is not simply whether schizophrenics differ from normals, but rather whether they differ from other psychiatric groups.

A fourth group of *S*s—college students—which can be assumed to be quite undisturbed in concept-forming behavior was included as a contrast sample. The samples tested were thus:

I. Schizophrenics clinically manifesting thinking disturbance (n = 18; mean age 35.5 years; mean education 11.06 years[3]).

II. Schizophrenics clinically not manifesting thinking disturbance (n = 20; mean age 40 years; mean education 11.16 years).

III. Non-schizophrenics, non-psychotic hospitalized psychiatric cases (n = 20; mean age 40.5 years; mean education 11.80 years).

IV. College students (volunteers from elementary psychology class at Stanford; n = 25; mean age 20.1 years; mean education 13.98 years).

For the hospitalized *S*s the diagnoses were as follows: For group I paranoid 11, catatonic 2, hebephrenic 3, mixed 2; for group II paranoid 15, mixed 5; and for group III anxiety 9, depressive 6, conversion 2, obsessive 1, alcoholism 1, nervous condition 1.

2. *Test Materials*

Eighty small cards (1¼ inches × ¾ inch) were prepared with the name of an object typed on each card. The objects were all non-emotional and the sorts of things with which an *S* would be familiar, e.g., stool, chair, cloth, leaf, finger, nose, shoe, cow, fish, cup, milk, mouth, tobacco, bear, engine, boat, rock, grass, and cigar.[4] The cards were placed in a tray (13

---

[3] Educational data were not available for all *S*s. Sample sizes on which years of education are based were, for Group I, 16; for Group II, 19; for Group III, 15; and for Group IV, 24.

[4] In addition to these object-names, the following were used:

Ford
Plymouth
Chevrolet
Dodge
table
bed
branch
plant
apple
peach
pear
seed
car
train
pie
wagon
truck
leg
eye
teeth
lettuce
tomato
highway
railroad
river
ocean
lake
elk
ice cream
cake
cart
knife
whale
sea weed
spoon
metal
elephant
ivory
sea shell
ash tray
cigarette
lawnmower
moon
box
pencil
ball
movie
ape
baseball
bell
coconut
cookie
pumpkin
pearls
flower
corn
blanket
room
tent
fruit
vegetable

inches × 11 inches) with the words up. The tray rested on a table at which *S* was seated. Instructions to *S* were:

> Here is a tray with a lot of cards on it. Each card has printed on it the name of some object. You are to arrange these cards in various groups. You can have as few or as many cards in each group as you wish. Perhaps I can best show you what I mean by looking for some cards that might go together. Here's *Ford,* and here's *Plymouth, and* they go together so I'll put them together (do it) and here's *Dodge,* so I'll put it there too. Do you get the idea now?
>
> All right, you go ahead: Use all of the cards. I want to emphasize that there are no right or wrong ways to do this. Everyone does it somewhat differently—I want to see how you do it. You will find that most of the cards can go in groups but if you feel that some of the cards don't belong, then you can just leave them in the tray. Let me know when you're through.

The objects had been so selected that there were a large number of different ways in which they could logically be grouped. *E* recorded the various groupings, the time required by *S*, the number of cards grouped, and *S*'s explanation (obtained after completion of the categorizing procedure) of the rationale for each grouping

In addition to the concept-forming task most of the patients *S*s were administered the WAIS (9) vocabulary test, and some of the *S*s were given McReynolds' (4, 5) Rorschach Concept Evaluation Technique (these procedures did not constitute a major part of the study).

3. *Procedure*

All *S*s were examined individually. Patient *S*s were seen as soon as possible after the judges' selections.

4. *Treatment of Data*

Each of the word groupings made by an *S*, along with his stated rationale for each grouping, was typed on a slip of paper. Altogether, for all *S*s, there were 1067 such groupings. The slips of paper bearing these groupings were put in random order, and two clinical psychologists, not otherwise involved in the study, judged each grouping in terms of the patient's stated rationales: these judgments were of course made without knowledge of which *S*s had made given card groupings. The task assumed by the judges was to assign to each card grouping a code based on a system devised by the authors. This system included 10 different codes intended to reflect empirical characteristics of the card-groupings rationales which could be given by the *Ss*. The authors did not attempt to devise a scoring system which would reflect any particular theoretical point of view; rather, their aim was simply to so characterize different ways of grouping cards together that it would be possible to code the particular ways utilized by different *Ss*. Table 1 summarizes the scoring system used. More detailed instructions were made available to the two judges. The *concept quality score* refers to values assigned by the two authors and the two judges as being representative of the over-all quality represented by the different codes. In making these assignments a high quality grouping was defined as one in which the cards were placed together according to an over-all conceptual principle which applied appropriately to all objects in the group, which was recognized and stated by *S*, which involved a major attribute of each of the objects, which was clear and communicable, and which was based on generally meaningful conceptual schemata rather than on personally idiosyncratic cues. It will be noted that low values represent a high level of concept-forming behavior. It is also to be emphasized that the judges, in coding the various groupings, considered *both* the nature of the cards grouped and the rationale given for the grouping.

TABLE 1
SYSTEM FOR CODING AND SCORING CARD GROUPINGS

Concept quality score	Code	Characterizations of groupings and patient's rationale for them
1	A	Rationale adequate and cards appropriate to it (apple, pear, peach . . ."fruit")
2	C	Rationale applicable to some, but not all cards included (elk, cow, elephant, ivory, tooth . . ."animals")
2	D	Rationale based on one card in group (animal, elk, cow, fox . . ."all animals")
3	B	Rationale vague and inadequate (lettuce, corn, tomato, pumpkin . . ."natural way to put them")
3	E	Rationale applicable to objects, but odd or awkward (poem, song, . . ."these are books")
4	H	Rationale consists of object words in phrase to show function (grass, lawnmower . . ."cut grass with a lawnmower")
4	I	Rationale consists of object words in chain association (boat, lake, fish, tent, bed . . ."in a boat on the lake, and caught a fish then went to the tent and went to bed")
5	F	Rationale, for objects grouped, is bizarre (napkin, diamond, saddle, ink, picture . . ."these are particles")
5	G	Rationale a combination of words (elk, tooth . . ."elk tooth")
	U	Unclassifiable, none of above

A *total conceptual quality score* was obtained for each *S* by computing a total score separately from the codings of each judge and averaging these two scores. For each judge the total score for a given *S* was obtained in the following manner. Each of the 80 cards was given the score represented by the code assigned to the group of which it was a part (e.g., all cards in groupings coded A were given a score of one, all cards in groupings coded F were given a score of five, and so on), and these scores were summed. This sum was divided by the number of cards scored (this was equal to the number grouped by *S* minus any which the judge had coded Unclassifiable). The final score for each *S* thus varied between one and five, and indicated directly the mean concept quality score for the cards grouped.

C. Results and Discussion

The reliability of the concept quality score, as estimated by correlating the scores computed separately for the two judges and correcting for attenuation, was .93. It is noted that this refers to reliability of scoring, not to reliability of the experimental procedure *per se*. No estimate of the latter was obtained, but this seems unnecessary in view of the positive findings to be reported below.

The major results are presented in Table 2. It may be noted first that the samples differed significantly on number of cards grouped. It follows, though it is not indicated in the table, that they also differed significantly on number of cards not grouped. Now, if the placement of cards in groups by an *S* is considered as being related to the process of assimilating percepts (6, 7)—an assumption which seems reasonable—then our findings lead us to believe that schizophrenics have more difficulty in assimilating common, everyday sorts of objects than do non-schizophrenics. The difficulty is somewhat greater for the thinking disturbance *S*s than for the non-thinking disturbance patients, but interestingly it is about the same for student and hospitalized non-schizophrenic psychiatric *S*s.

The samples do not differ significantly on numbers of groups formed; they do, however, differ on *rate* of categorizing the cards, with the thinking disturbance *S*s being remarkably slower than the other three samples. Now, the rate at which one assimilates incoming percepts has been put forward elsewhere (6, 7) as a major personality variable, and it has been suggested that a low rate may be implicated in certain schizophrenic symptomatology similar to that induced by sensory deprivation. It is therefore noteworthy that the *S*s with the lowest rate of categorizing the stimuli is the sample which has most pronounced symptomatology of the kind induced by sensory deprivation (i.e., thinking disturbance).

The samples differed significantly on the concept quality score. The differences were in the order that any clinician would predict, but in addition to the over-all findings there are several comparisons deserving of comment. First, it is noted that the two schizophrenic samples have quite similar scores (they are not significantly different). This suggests that aberration in concept-forming behavior is not limited to schizophrenics with clinically manifest thinking disorder, but that it also characterizes—though to a lesser degree—schizophrenics not showing much clinical disorganization. The impression that poor quality conceptual behavior is particularly characteristic of schizophrenia is further supported by the findings for the other two samples. Both of these differed significantly from the schizophrenic *S*s, but they did not differ significantly from each other. The fact that the hospitalized non-psychotic psychiatric patients did not form significantly poorer quality concepts than the students may of course reflect an inadequacy in the experimental procedure, but is seems a more likely interpretation—in view of the over-all findings—to suppose that their particular disorders were not of the kind or severity which would infringe markedly upon thinking processes.

TABLE 2

COMPARISONS OF MEANS ON FOUR INDICES OF CONCEPT-FORMING BEHAVIOR

Sample	Means no. cards grouped	Mean no. groups	Mean sec. used per card*	Mean concept quality score
Schiz. with thinking disturbance (n = 18)	52.50	12.06	24.76 (n = 14)	2.32
Schiz. without thinking disturb. (n = 20)	65.90	13.25	14.52 (n = 17)	2.23
Hospitalized non-schizophrenics (n = 20)	71.20	12.85	13.69 (n = 11)	1.85
College students (n = 25)	73.96	13.12	12.22 (n = 24)	1.73
df	3/79	3/79	3/62	3/79
F	8.22	—	3.09	5.30
p	.005	n.s.	.05	.005

* Data not available on all *S*s; see values within parentheses.

It was possible to evaluate the test protocols in terms of the cards assigned by the judges to each of the different codes. It had been hoped that it would be possible to compare the different samples separately on each of the different codes. Thus, it was anticipated that schizophrenic *S*s would group the cards according to code F (bizarre) more than would controls and students, that students would group the cards according to code A more than the other samples, and so on. This part of the study, however, was not very successful—for two related reasons. First, it turned out that code A was used far more than any other code (even for schizophrenic *S*s). Percentages of cards given each code (mean for the two judges, for all samples combined) were: A, 40; B, 10; C, 8; D, 7; E, 8; F, 2; G, 1; H, 2; I, 3; and U, 1. (These total 82 per cent; the other 18 per cent of the cards were not categorized by the *S*s, and hence not coded by the judges). Second, reliabilities of the judges' ratings, while quite adequate for the test as a whole, were not outstanding for the separate codes, which may be thought of as items in a test. The correlation coefficients between the number of cards for each *S* assigned a given code by the two judges ranged from .80 (code A) to .10 (code I), with the mean of the r's being .51. As it turned out, the number of cards coded A was the only one reaching significance ($p = .005$) for the ratings made by both judges when the four samples of *S*s were compared. The fact that the samples did not differ significantly on the other codes appears to be more indicative of a shortcoming in the present study than of a true lack of difference in ways of categorizing percepts among the *S*s.

The hospital samples did not differ appreciably on education, and while Group I was somewhat younger than Groups II and III this cannot be considered a biasing factor since it was Group I which was poorest in concept-forming behavior, and age—if it has an effect at all—would be expected to work in the opposite direction. WAIS vocabulary scores were available on 14 *S*s from Group I, 18 from Group II and 19 from Group III. The respective means (raw scores) were 28.50, 43.50, and 58.63. These are significantly different at the .005 level. For the 51 *S*s just referred to the correlation between vocabulary and concept quality score was —.46, which is significant at the .01 level. If we think of the vocabulary results as measures of intelligence then the findings suggest the possibility that the differences among the samples on concept quality scores reflected differences in intelligence rather than differences in underlying pathology. While this possibility cannot be completely discounted, it seems more likely that the vocabulary scores are themselves indicative of lowered functioning intelligence as a result of pathology. It is notoriously difficult to obtain adequate measures of premorbid intelligence from psychotics as overtly and seriously ill as were those used in this study. Further both vocabulary and concept formation are aspects of general intelligence and so would naturally be positively related.

A final result concerned the comparison of the samples on McReynolds' (4, 5) Rorschach Concept Evaluation Technique (CET). The inclusion of this test in the study was not a part of the original plan, but protocols were available on nine *S*s in Group I, 17 in Group II, and 25 in Group IV. The two major variables on this test are labeled J and V. Of these two V has previously been held to reflect disturbance in thinking. It is significant, therefore, that in the present study the samples compared did not differ significantly on J, but did differ on V ($df = 50$; $F = 9.79$; $p = < .001$). The means (T-scores) were: for Group I, 34.33; for Group II, 45.41; and for Group IV, 46.92.

We do not believe that the exact procedure used in the present study is suitable for use as a clinical diagnostic tool. Our aim in the present research was primarily theoretical, and much additional work would be required in order to develop the procedure into a standardized clinical instrument. In particular, it is our belief that the scoring system used here could be made less cumbersome and more precise.

However, it does appear that the card categorizing technique used in this study could be adapted to additional research questions, especially since it appears to be an analogue of the process whereby a person assimilates his percepts into his apperceptive mass. For example, it would be possible to include cards representing emotional and non-emotional stimuli, in order to determine if there is a difference in the speed and adequacy with which these two classes are categorized. Further, one could investigate the extent to which adequacy of categorization is related to particular problem areas of *S*s.

D. SUMMARY

This study compared concept-forming behavior in (1) schizophrenics having clinically manifest thinking disturbances; (2) schizophrenics not having clinically manifest thinking disturbances; (3) hospitalized non-psychotic psychiatric patients; and (4) college students. The design permitted a wide sampling of the *S*s' concept-forming behavior. It was found that the groups of *S*s differed in quality of concepts formed, in the reverse order in which the groups are listed above. However, the two groups of schizophrenics did not differ significantly from each other, nor did the two non-schizophrenic groups. This implies that, within the range of *S*s sampled in this study, poor concept formation is particularly characteristic of schizophrenia. It is noted, however, that most of the concepts formed even by the schizophrenic *S*s were rated as adequate, i.e., that inadequate concept formation is by no means characteristic of all the thinking of schizophrenics. It is believed that the general technique used in this study can be usefully applied to other research questions.

REFERENCES

1. BRUNER, J., GOODNOW, J., & AUSTIN, G. A Study of Thinking. New York: Wiley, 1956.
2. CHAPMAN, L. J., & TAYLOR, J. Breadth of deviate concepts used by schizophrenics. *J. Abnorm. & Soc. Psychol.*, 1957, **54**, 118-123.
3. McGAUGHRAN, L. S., & MORAN, L. J. "Conceptual level" vs. "conceptual area" analysis of object-sorting behavior of schizophrenic and nonpsychiatric groups. *J. Abnorm. & Soc. Psychol.*, 1956, **52**, 43-51.
4. McREYNOLDS, P. Perception of Rorschach concepts as related to personality deviations. *J. Abnorm. & Soc. Psychol.*, 1951, **46**, 131-141.
5. ———. The Rorschach concept evaluation technique. *J. Proj. Tech.*, 1954, **18**, 60-72.
6. ———. A restricted conceptualization of human anxiety and motivation. *Psychol. Reports* Monogr. Suppl. No. 6, 1956, 293-312.
7. ———. Anxiety, perception, and schizophrenia. *In* Jackson, D. (*Ed.*), The *Etiology of Schizophrenia.* New York: Basic Books, 1960.
8. PIAGET, J. The Origins of Intelligence in Children. New York: Internat. Univ. Press, 1952.
9. WECHSLER, D. Wechsler Adult Intelligence Scale. New York Psychol. Corp., 1955.
10. WINDER, C. L. Some psychological studies of schizophrenics. *In* Jackson, D. (*Ed.*), *The Etiology of Schizophrenia.* New York: Basic Books, 1960.

VA Hospital
3801 Junipero Serra Boulevard
Palo Alto, California

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 379-402.

REINFORCEMENT AND PERSONALITY FACTORS IN VERBAL CONDITIONING*

Department of Psychology, Indiana University

ARNOLD BINDER AND PHYLLIS SALOP

A. INTRODUCTION

The similarities between the methodology of the operant conditioning of verbal behavior and the process of psychotherapy and similar interpersonal activities have been emphasized by a number of writers. In fact many argue that verbal conditioning provides an experimental paradigm of the interpersonal process, in general, or the psychotherapeutic process, specifically. To the extent that this position is valid, it is important to investigate by means of verbal conditioning experiments those variables which are important in the interpersonal-psychotherapeutic area.

Two such variables are the personality of the person whose verbal behavior is reinforced and the type of stimuli used as reinforcers. Both of these variables were investigated in the present experiment.

Krasner (1958) has provided a very thorough review of the literature on operant verbal conditioning. The general conclusions based on this survey are: (*a*) significant increases in the verbal responses reinforced were obtained with such reinforcers as "good," "mmm-hmm," a head nod, and a light; (*b*) significant decreases in the category of response reinforced were obtained with such experimenter responses as "huh-uh" and a head shake; (*c*) effective response classes included plural nouns, "hostile" verbs, various pronoun groupings, and statements of opinion; (*d*) the conditioned responses are very resistant to extinction; (*e*) learning was shown even when the *S*s failed to answer such direct questions as, "When and why did I say 'good'?" at the conclusion of the experiments. Of particular interest for the present research are the studies which related conditionability to personality variables. The principal findings among these studies are: (*a*) higher scores on tests of anxiety like the Taylor Manifest Anxiety Scale (MAS) tend to be associated with greater amounts of verbal conditioning (although this was not supported in more recent research), (*b*) the personality of the examiner is an important determinant of rate of verbal conditioning, even when the experimental

---

* Received in the Editorial Office on July 11, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

procedure is mechanical and inflexible, (*c*) *S*s with the personality characteristics specified by such constructs as "lack of protection" and "compliant" performed at higher levels of verbal conditioning, while those with characteristics like "defensiveness" and "noncompliant" performed at lower levels, and (*d*) stress inducing instructions and nonstress instructions produced no differential conditioning.

Although several investigations dealing with personality variables and verbal conditioning have been published subsequent to Krasner's review (these will be discussed later), there is still very little systematic evidence on the relationship between personality patterns, on the one hand, and the rate of verbal conditioning, on the other. Furthermore, there is almost no information available on the interactive effects of personality variables and type of reinforcing stimulus upon conditioning rate. To help remedy these deficiencies, in the present experiment the personalities of all *S*s were evaluated by means of the Minnesota Multiphasic Personality Inventory (MMPI) and the conditions of reinforcement were varied for different groups of *S*s. Thus, one group was reinforced by *E*'s saying "good" after each critical response, another group was given electric shock, while a third group was given no reinforcement.

B. METHOD

1. *Subjects*

The *S*s were all males enrolled in a section of introductory psychology at Indiana University. All *S*s were given the MMPI in large groups prior to participation in the experiment. The *S*s were assigned to one of three reinforcement groups on a haphazard basis. The same female experimenter (*E*) conducted the verbal conditioning sessions for all *S*s.

2. *Materials and Apparatus*

The stimulus materials shown to the *S*s consisted of 150 3″ × 5″ white, unlined cards with four pronouns and two verbs typed on each card. The pronouns were typed on a single line a little above the center-line across the 5″ dimension, and the verbs were typed in a single line just below this line. The pronouns *you, we, they, I* appeared on every card; however, their order was randomized over cards. The two verbs on each card differed in content and tense; one was in the past tense while the other was in the present tense. Moreover, a given verb appeared on only one of the cards.

The total of 300 verbs used in the experiment were selected from a much larger listing of verbs on the basis of ease of pronunciation, simplicity, lack of emotional connotations, and general appropriateness. These verbs were divided into 150 arbitrary pairs and each pair was typed on only one card. The particular verb in each pair put in the past tense and that put in the present tense was decided randomly. The verb pairs actually used may be seen in Table 1.

The complete set of 150 cards was grouped haphazardly into 15 stacks of 10 cards each. One stack of 10 cards then provided the stimulus materials for one learning block.

The apparatus used to present the cards consisted of a face board (10½" × 18½") and two side boards (10½" × 19"). The face board was placed perpendicular to the line of sight of the *S*s, and one end of each side board was attached to an end of the face board. The side boards then extended back toward *E*. The face board supported a holder for the 3" × 5" cards, and all three boards served to conceal the manipulations of the cards by *E*.

There was a 3" × 5" card in the holder visible to each *S* as he sat down before the presentation apparatus. The card contained the same four pronouns as those on the other cards and two verbs which did not occur elsewhere in the experiment. The *S* was then told that he would be shown a series of cards like the one before him, containing four words in the upper part and two words in the lower part. He was instructed further to make up a sentence using one of the four pronouns and one of the two verbs. It was emphasized that he was to respond as rapidly as possible after each new card was presented, without analyzing the sentence or trying to censor or change it in any way. It was pointed out that each sentence could be of any length or structure so long as it contained one of the pronouns and one of the verbs.

After *S* indicated that the instructions were understood, *E* presented the first experimental card. Each subsequent card was presented immediately after *S* responded with the sentence appropriate for the card at hand. Although the total of 150 cards was divided into 15 stacks of 10 cards each, there was no change in pattern of presentation when the first card of a new group followed the last card of the preceding group.

Prior to the arrival of a new *S*, *E* shuffled each of the 15 stacks and haphazardly shifted the ordering of the stacks. For all *S*s no reinforcement followed the sentences given in response to the first stack of cards. Beginning with stack two and extending through stack nine, *E* responded "good" immediately after the *S*s in one of the three groups used a sentence containing the verb in the past tense. The particular choice of pronoun had no effect

TABLE 1
VERB PAIRS PRESENTED TO *Ss*

fall	indicated	jump	ate	originated	fill
enclose	obeyed	matched	illustrate	spill	baked
enter	boiled	admired	deserve	smiled	knock
earn	nominated	comment	photographed	told	pour
abandon	verified	contributed	arrange	stand	marked
clean	recovered	declared	predict	marched	adjust
hinted	insure	opened	identify	elevate	recalled
invited	create	settled	know	deposited	check
overlooked	finish	slide	caught	imagine	enjoyed
maintain	depended	worked	place	invent	concluded
measure	injected	labor	dreamed	listed	stop
inscribe	played	manipulate	informed	grew	throw
emphasized	prolong	guarded	pause	encouraged	recommend
discover	watched	folded	borrow	examined	move
decided	react	came	ponder	hide	built
deviated	call	draw	eliminated	introduce	infected
construct	mixed	organize	completed	mailed	install
held	make	remarked	persist	noticed	grant
communicated	paint	weigh	collected	inspected	imitate
turned	cancel	walk	established	traveled	decorate
inspired	hear	observe	planned	listened	point
furnish	desired	lifted	drop	liberate	separated
fix	departed	improve	gained	follow	wondered
expected	magnify	greet	demonstrated	emptied	gather
directed	mention	fed	investigate	bounced	reach
meet	interpreted	explained	order	blow	elected
prefer	hoped	broke	combine	answered	get
talked	look	determined	encircle	continue	paid
passed	name	retired	foresee	derived	include
accept	gave	practice	danced	occupy	needed
altered	have	pull	discussed	join	influenced
learn	ran	kept	compose	divide	bent
fished	glance	grasp	transferred	regulated	repeat
flip	carried	advise	cheered	permit	registered
counted	formulate	act	breathed	surrounded	return
close	hired	declare	polished	relieved	entertain
compared	distribute	chose	interfere	omit	hurried
realized	replace	promised	taste	noticed	find
respond	left	sailed	protect	wanted	discriminate
washed	inherit	produce	went	spoke	employ
receive	bought	lock	descended	ignore	participated
justified	change	live	saw	guided	happen
handle	agreed	insert	corrected	forget	defended
impressed	promote	gaze	drove	appear	commended
diagnosed	pronounce	devise	picked	allowed	blot
numbered	lead	deliver	added	manage	acknowledged
save	provided	send	loaded	sell	tried
whisper	applied	recorded	remember	stepped	begin
sprinkled	help	succeed	wrote	swam	hinder
thought	file	recognize	admitted	rode	slip

in determining whether reinforcement did or did not occur. There were 50 *S*s in this group, which we may conveniently designate as Group G.

Similarly, another group of *S*s was given a mild electric shock whenever a sentence containing the verb in the past tense was used, between the second and ninth stacks inclusively. There were 39 *S*s in this group, designated Group S. The intensity of the shock administered was adjusted in accordance with the verbal report of each *S*. Beginning with a level of two milliamperes the shock was increased in steps of two milliamperes, to a maximum of 10 milliamperes, until the *S* reported that the shock was just painful or the maximum was reached. With the consent of *S*, the shock was then increased by one milliampere, unless the maximum had already been reached.

Reinforcement was stopped after stack nine for Groups G and S, and no further reinforcement was given through blocks 10 to 15. The third group of *S*s, Group C, served as controls and received no reinforcement throughout the experiment. There were 19 *S*s in Group C.

After the 150 trials had been completed, each *S* in the reinforcement groups was asked questions to determine the facility with which he could verbalize the response class and reinforcement contingency. He was first asked if he knew what the experiment was about. Then, if he was in Group G he was asked, "Do you have any idea of when or why I was saying 'good'?," while if he was in Group S he was asked, "Do you have any idea of when or why you were being shocked?."

C. RESULTS

Each stack of cards, although not differentiated for the *S*s, was considered a learning block in analyzing the data. The learning score for each *S* in a given block consists of the number of sentences involving the use of the past tense (rather than present tense) verbs in response to the cards in the relevant stack. Figure 1 shows the learning curves for the three groups in terms of mean number of sentences per block involving past tense verbs. Only the data for those *S*s in Groups G and S who did not verbalize the relationship between their behavior and reinforcement in response to the specific questions are included in the figure, as well as in all the analyses to follow. Since eleven of the *S*s in Group G and two of the *S*s in Group S did verbalize the relationship, the numbers of *S*s available for the analyses in these groups were 39 and 37, respectively.

Initial comparisons of learning trends involved Groups G and C, on the one hand, and Groups S and C, on the other. The statistical procedure for these comparisons was the test for comparing group trends developed by Alexander (1946). The results of these tests are summarized in Table 2

TABLE 2
SUMMARY OF THE TREND TESTS COMPARING REINFORCED GROUPS WITH CONTROL GROUP (ACQUISITION)

	Group G vs. Group C				Group S vs. Group C			
Source	SS	df	MS	F	SS	df	MS	F
Individual Deviations from estimation (error)	729.91	392	1.86		824.52	378	2.18	
Between individual slopes	163.98	56	2.93		155.00	54	2.87	
Between individual means	1071.89	56	19.14		904.31	54	16.74	
Group deviations from estimation	14.87	7	2.12		19.51	7	2.79	
Between group slopes	7.00	1	7.00	3.75	.32	1	.32	.14
Between group means	35.85	1	35.85	19.25**	6.04	1	6.04	2.76
Over-all deviations from linearity	15.20	7	2.17	1.16	11.53	7	1.64	.75
Over-all slope	9.72	1	9.72	5.22*	.01	1	.01	.00
Total	2048.41	521			1921.23	503		

* $P \leqslant .05$
** $P \leqslant .01$

TABLE 3

SUMMARY OF THE TREND TESTS COMPARING REINFORCED GROUPS WITH CONTROL GROUP (EXTINCTION)

Source	Group G vs. Group C SS	df	MS	F	Group S vs. Group C SS	df	MS	F
Individual deviations from estimation (error)	405.48	224	1.81		505.21	216	2.34	
Between individual slopes	123.58	56	2.21		159.81	54	2.96	
Between individual means	842.63	56	15.05		981.82	54	18.18	
Group deviations from estimation	1.08	4	.27		1.62	4	.40	
Between group slopes	.62	1	.62	.34	1.23	1	1.23	.52
Between group means	19.46	1	19.46	10.75**	33.44	1	33.44	14.29**
Over-all deviations from linearity	3.93	4	.98	.54	2.30	4	.58	.24
Over-all slope	.15	1	.15	.08	.00	1	.00	.00
Total	1396.93	347			1685.43	335		

** $P \leqslant .01$

for the acquisition trials (blocks one to nine), and in Table 3 for the extinction trials (blocks 10 to 15). The differences between group means are significant ($P \leqslant .01$) for both comparisons on extinction and for the comparison between Groups G and C on acquisition; the group means do not differ significantly between Groups S and C on acquisition. The differences between group slopes (that is, linear trends) are not significant in all four comparisons, but the *F* resulting from the comparison involving Groups G and C on acquisition is only 11 hundredths short of significance at the .05 level (the estimated significance level is .054).

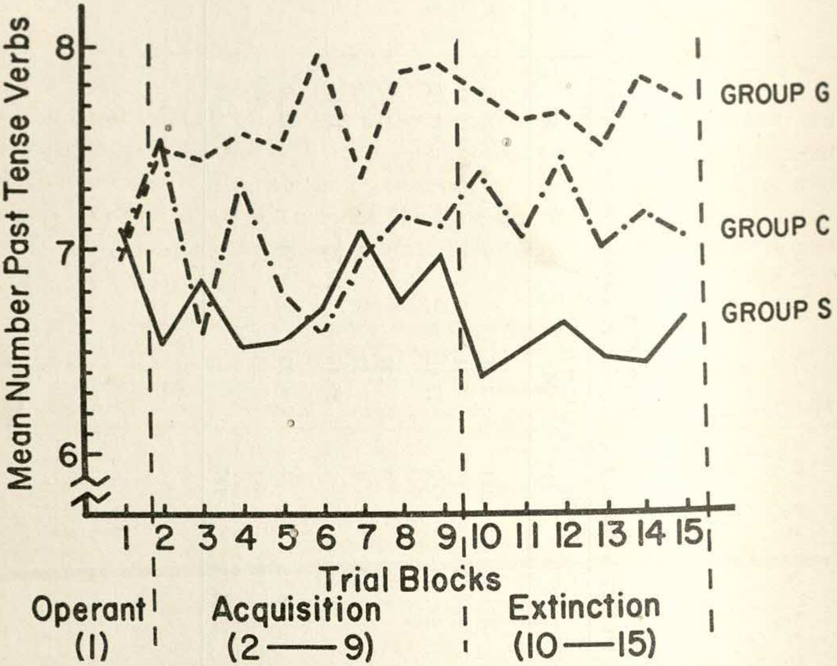


FIGURE 1

Learning Curves for the Three Reinforcement Groups

Because of this close result, there was some interest in further tests of group slope by means of analyses of the groups taken separately in the manner described by Alexander (1946). These results are summarized in Table 4. The linear learning trend for acquisition is significant ($P \leqslant .01$) for Group G, but not for Groups S and C. None of the trends during extinction are significant.

For the further analyses, the median score on each of the 10 most commonly used MMPI clinical scales plus the median of the distribution of the Hy-Pt scores were computed for each of the groups. The learning trend of those individuals having MMPI scores above the median on a given scale was then compared with the trend of individuals having MMPI scores below the median on that scale by Alexander's test. This was done separately for each group.[1] These results are summarized in Table 5. Only the final *F* values for Between Group Slopes and Between Group Means are shown because of the almost overwhelming array of results.

TABLE 4
F-VALUES FROM TREND TESTS FOR SEPARATE GROUPS

Group	Acquisition	Extinction
G	8.65**	.00
S	.08	.20
C	.07	.31

** $P \leqslant .01$

TABLE 5
F-VALUES FOR COMPARISONS BETWEEN HIGH AND LOW SCORERS ON MMPI SCALES (ACQUISITION)

	Group G		Group S	
MMPI scale	Between slopes	Between means	Between slopes	Between means
Hs	.12	.46	2.16	.31
D	.58	.00	.02	14.67**
Hy	.05	22.48**	.33	5.63*
Pd	.28	5.44*	1.89	10.36**
Mf	8.39**	30.17**	.35	.46
Pa	.05	3.56	.03	9.62**
Pt	2.16	4.41*	.36	12.34**
Sc	.04	.25	1.50	1.89
Ma	4.24*	.53	1.64	2.37
Si	1.65	7.14**	.28	14.91**
Hy-Pt	.40	.12	.83	14.24**

* $P \leqslant .05$
** $P \leqslant .01$

The only significant differences between group slopes among the acquisition comparisons were those involving the Mf and Ma scales for Group G. Significant differences between group means were obtained on Hy, Pd, Mf, Pt, and Si for Group G; and on D, Hy, Pd, Pa, Pt, Si, and Hy-Pt for

---

[1] Running all these tests would have been extremely difficult and time-consuming without the aid of an electronic computer. The programming for the IBM650 of Alexander's test for trend and for comparing group trends are available from Indiana University.

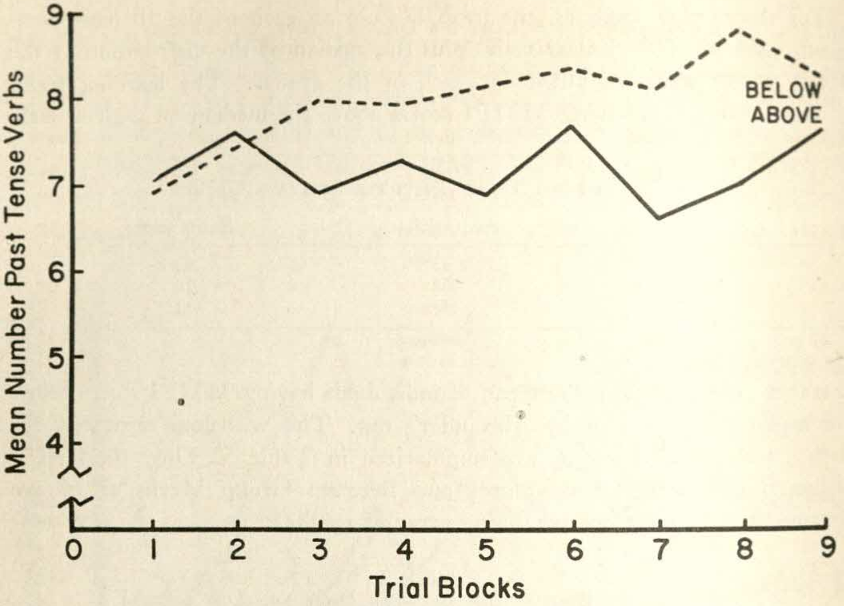


FIGURE 2
LEARNING CURVES COMPARING GROUPS ABOVE AND BELOW MEDIAN ON MF (GROUP G—ACQUISITION)

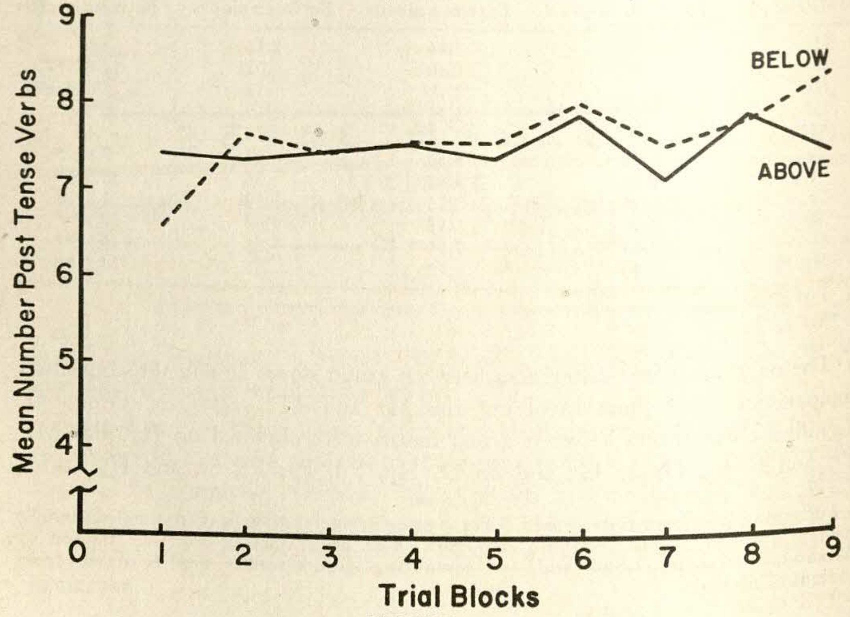


FIGURE 3
LEARNING CURVES COMPARING GROUPS ABOVE AND BELOW MEDIAN ON MA (GROUP G—ACQUISITION)

TABLE 6
COMPARISON OF MEANS FOR GROUPS ABOVE AND BELOW MEDIAN ON MMPI (ACQUISITION)

Group	Scale	Group above or below median	Block 1 (Operant)	Mean over blocks 2 to 9
G	Hy	Above	7.32	7.99
		Below	6.65	7.29
G	Pd	Above	7.44	7.92
		Below	6.56	7.65
G	Mf	Above	7.05	7.14
		Below	6.90	8.10
G	Pt	Above	7.93	7.51
		Below	6.78	8.01
G	Si	Above	7.38	7.38
		Below	6.91	7.85
S	D	Above	6.89	7.13
		Below	7.26	6.38
S	Hy	Above	6.56	6.58
		Below	7.58	6.89
S	Pd	Above	7.22	7.03
		Below	6.95	6.47
S	Pa	Above	7.10	7.02
		Below	7.06	6.41
S	Pt	Above	6.82	7.11
		Below	7.24	6.39
S	Si	Above	6.89	7.06
		Below	7.17	6.30
S	Hy-Pt	Above	7.19	6.47
		Below	7.53	7.15

TABLE 7
F-VALUES FOR COMPARISONS BETWEEN HIGH AND LOW SCORERS ON MMPI SCALES (EXTINCTION)

	Group G		Group S	
MMPI scale	Between slopes	Between means	Between slopes	Between means
Hs	7.18**	.83	11.57**	.70
D	10.34**	.85	.12	32.01**
Hy	3.69	4.06*	2.16	17.32**
Pd	3.77	.55	.09	8.99**
Mf	.46	39.32**	3.23	.91
Pa	1.02	.96	1.08	4.51*
Pt	2.85	4.62*	.07	10.85**
Sc	5.40*	1.03	3.27	9.02**
Ma	2.33	17.50**	.60	4.40*
Si	4.79*	11.39**	3.32	4.69*
Hy-Pt	2.40	.01	3.44	15.61**

* $P \leqslant .05$
** $P \leqslant .01$

Group S. The learning curves for the groups above and below the median on the two scales containing significant differences in slope are shown in Figures 2 and 3. To make the most meaningful comparison between the high and low MMPI groups on the scales containing significant differences in group mean, only the mean for block one and the over-all mean for blocks two to nine are shown. These are in Table 6.

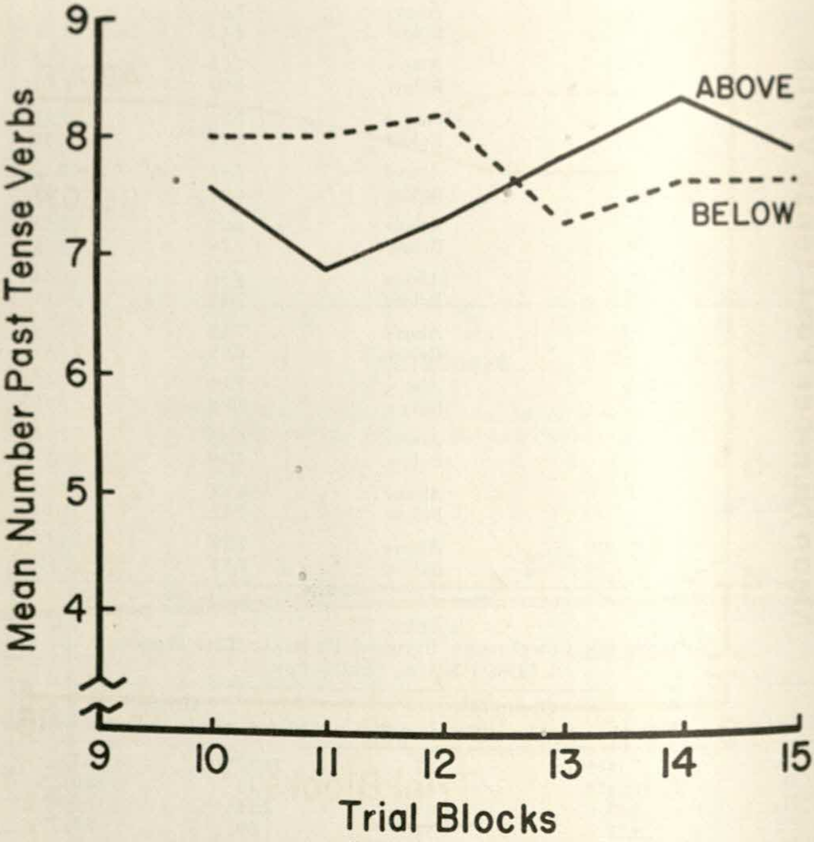


FIGURE 4
LEARNING CURVES COMPARING GROUPS ABOVE AND BELOW MEDIAN ON Hs (GROUP G—EXTINCTION)

On the extinction trials (see Table 7), significant differences in group slope were found on Hs, D, Sc, and Si for Group G; and on Hs for Group S. Similarly, significant differences between group means were found on

Hy, Mf, Pt, Ma, and Si for Group G; and on D, Hy, Pd, Pa, Pt, Sc, Ma, Si, and Hy-Pt for Group S. Figures 4, 5, 6, 7, and 8 show the extinction curves for the high and low MMPI groups differing significantly in slope. And Table 8 contains over-all means for blocks eight and nine, on the one hand, and blocks 10-15, on the other, for the high and low MMPI groups having significant mean differences on extinction.

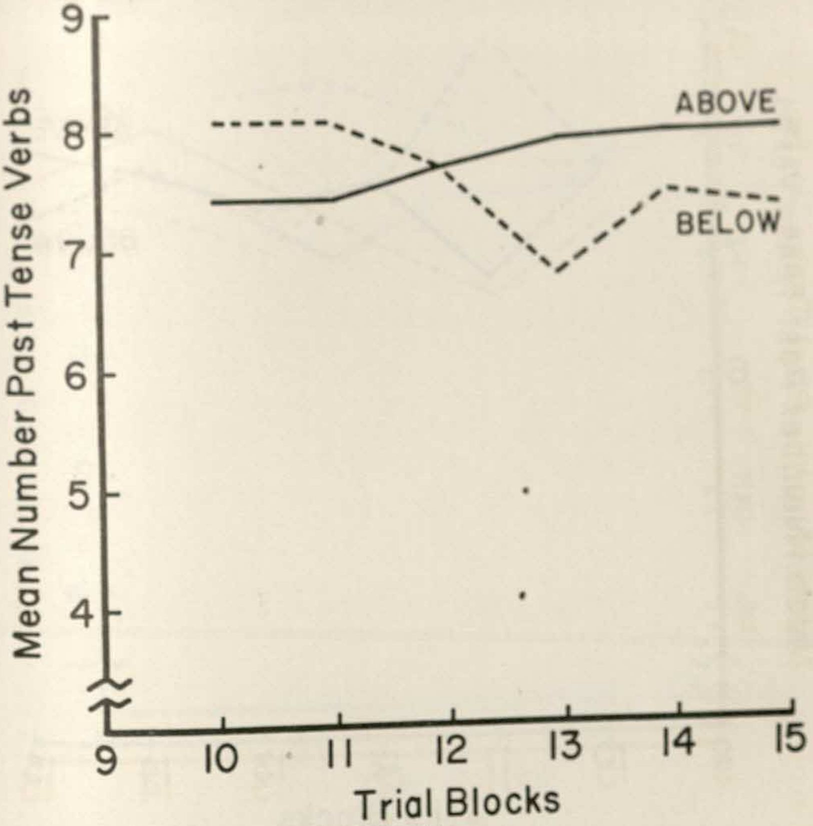


FIGURE 5
LEARNING CURVES COMPARING GROUPS ABOVE AND BELOW MEDIAN ON D (GROUP G—EXTINCTION)

D. DISCUSSION

In accordance with the customary practice in this type of experiment, only those *S*s who did not indicate the relationship between their responses and the reinforcement were included in the preceding analyses. Previous investigators have distinguished between verbalizers and nonverbalizers by such terms as aware–not aware, insightful-noninsightful, and conscious-unconscious, with inferential accompaniments comparable to the level of the constructs used. A number of papers (especially recently) have challenged the procedure of determining awareness or insight on the basis of the responses to a restricted set of questions. The arguments have been, essentially, that the questions may not be thorough enough to determine actual awareness or that the verbalizations of the *S* may be invalid in not representing the true state

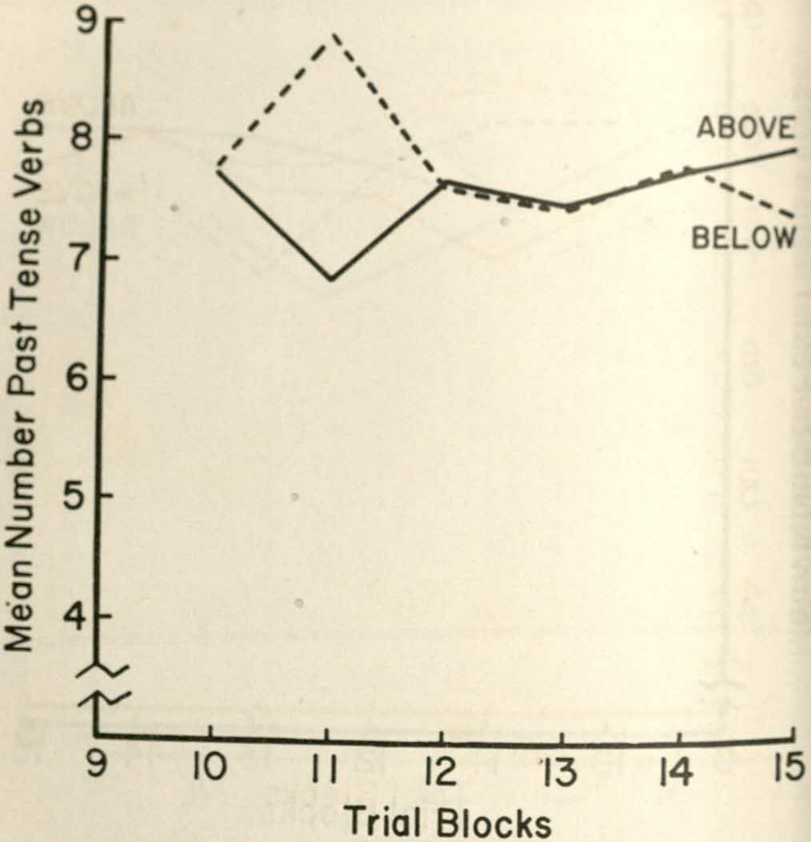


FIGURE 6
LEARNING CURVES COMPARING GROUPS ABOVE AND BELOW MEDIAN ON Sc (GROUP G—EXTINCTION)

of affairs. Among the advocates of this position are Krasner, Weiss, and Ullman (1959), Tatz (1956), Levin (1961), and Krieckhaus and Eriksen (1960).

Awareness is obviously a construct based upon specific observations of behavior. For the most part these observations are verbal, but mannerisms, gestures, and other incidental motor phenomena may be used in anchoring it. The usefulness of a given method of anchoring the concept depends upon its role in theory, and not on any metaphysical conceptions of what aware and unconscious "really are." Awareness may be defined in terms of a few general

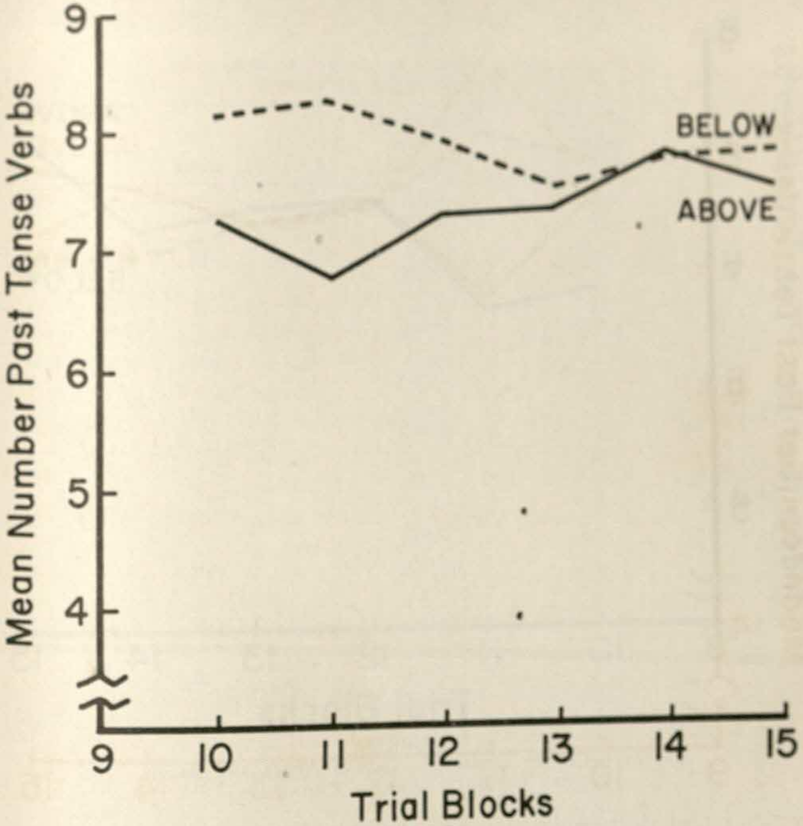


FIGURE 7
LEARNING CURVES COMPARING GROUPS ABOVE AND BELOW MEDIAN ON SI (GROUP G—EXTINCTION)

questions (Greenspoon, 1955; Taffel, 1955; Binder et al., 1957), a more probing array of questions (Tatz, 1956; Krieckhaus & Eriksen, 1960), a searching interview (Levin, 1961; Krasner et al., 1959), or upon the incidental post-experimental gossip of *S*s (Krieckhaus & Eriksen, 1960; Krasner, 1958), but in any case the defining process is a function of the theoretical position for which the given research is supplying further information, as well as the stage of development of the theory at the particular time.

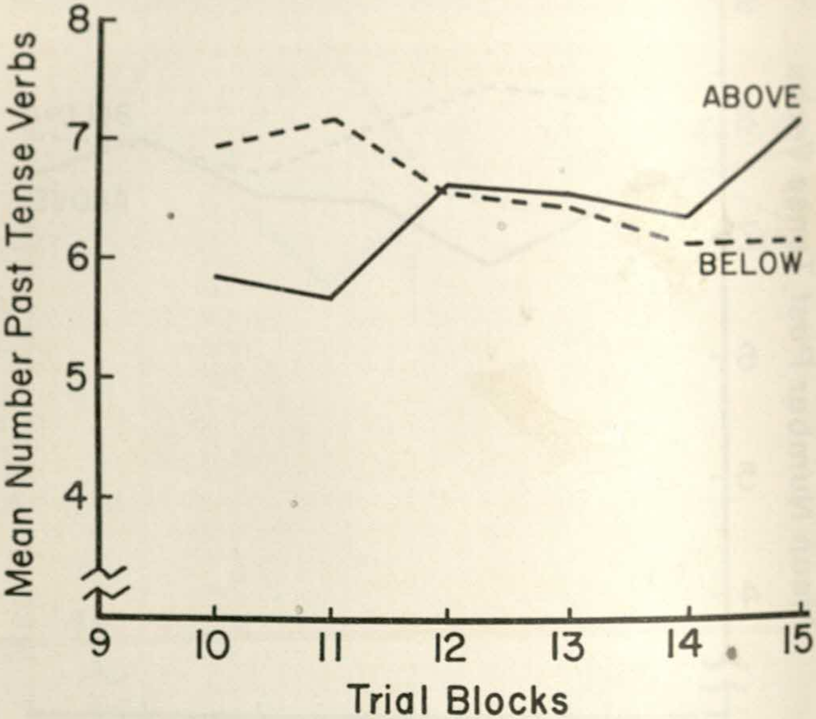


FIGURE 8
Learning Curves Comparing Groups Above and Below Median on Hs (Group S—Extinction)

It is difficult to see what the scientific meaning of the terms aware and awareness is to investigators who claim, "Clearly, it is impossible to *know* what *S* was aware of during the conditioning task" (Levin, 1961, p. 74) or "It is possible that learning, particularly responses such as verbal behavior which are so intimately associated with awareness, is an event which requires enough concentration of psychic energy to assure at least, its temporary establishment in awareness" (Krieckhaus & Eriksen, 1960). It is quite possible that one could establish very interesting and important empirical laws purely on the basis of defining awareness and unawareness in verbal conditioning on the basis of the responses of "Yes" and "No," respectively, to the question, "Do you know why I said 'good' during the experiment?" There is no doubt that the quantity of *S*s in the category unaware would diminish as the list of questions increased in length and specificity. The argument that the *S*s eliminated from the unaware group as defined by a given set of questions on the basis of one more specific question or a wink or a bit of

TABLE 8
COMPARISON OF MEANS FOR GROUPS ABOVE AND BELOW MEDIAN ON MMPI (EXTINCTION)

Group	Scale	Group above or below median	Mean over blocks 8 and 9	Mean over blocks 10 to 15
G	Hy	Above Below	8.21 7.58	7.85 7.52
G	Mf	Above Below	7.18 8.50	7.15 8.18
G	Pt	Above Below	7.77 8.20	7.52 7.91
G	Ma	Above Below	7.68 8.11	7.34 8.04
G	Si	Above Below	7.64 8.10	7.38 7.94
S	D	Above Below	6.94 6.76	7.10 5.95
S	Hy	Above Below	6.47 7.21	6.07 6.93
S	Pd	Above Below	6.89 6.82	6.82 6.21
S	Pa	Above Below	6.58 6.59	6.71 6.27
S	Pt	Above Below	6.79 7.00	6.82 6.11
S	Sc	Above Below	6.53 7.31	6.21 6.82
S	Ma	Above Below	6.72 7.19	6.30 6.77
S	Si	Above Below	7.00 6.58	6.67 6.22
S	Hy-Pt	Above Below	6.59 7.27	6.24 7.08

gossip were not really unaware in the first place carries a number of theoretical and philosophical risks. It is entirely conceivable that, within a certain theory, one indicator of awareness will be the very fact of increasing in the usage of a certain class of words when given a subtle reinforcer. Then clearly one could not possibly have verbal learning without awareness.

That is not to say that a given set of questions regarding awareness is as useful as any other set. It may be, for example, that a given set of questions does not rule out a simpler and well-established explanation for the obtained results. Included in this category are the factors of unreliability of evaluating the answers of the *S*s, the motivation of the *S*s in answering, the possible effects of forgetting, and deliberate attempts to please the examiner by giving certain responses. But these are quite different from the arguments to the effect that unawareness is equivalent to insensitivity to all cues. It would be folly to argue that *S*s classified as unaware have no partial solution or partial cues as to the conditioning process as determined by some pattern of examination or line of inquiry (including, perhaps, torture).

The justification for the particular questioning of the present experiment rests entirely upon the potentiality of establishing interesting empirical relationships. What are the differences in personality between individuals who do and those who do not condition, when only those are included who do not verbalize the reinforcement contingency in response to simple, direct questioning? It would seem reasonable to suppose at this rather impoverished stage in the accumulation of knowledge about verbal conditioning that somewhat different results would be obtained with a different line of questioning. But different results are not necessarily more important from the viewpoint of theory construction.

The findings of the present research are obviously varied and complex. A glance at Figure 1 shows a rather clear separation of the learning curves for the three groups, but the Group C and Group S *S*s show differences only after the seventh trial block. The acquisition slope for Group G is significant while the acquisition slopes for Groups C and S are not, but the difference in slopes between Groups G and C is just short of significance. In this connection it should be noted that the Alexander test is sensitive only to linear increases and that Group G shows a pattern of immediate, rapid increase of past tense verbs followed by slight increases thereafter. Considering that Group C gave slightly more past tense verbs than Group G during the operant block, the significant difference in means between the groups makes it aparent that a verbal conditioning effect was achieved with Group G. On the other hand there is no evidence of learning during the acquisition trials of Group S.

Although there are very distinct differences among the three groups in over-all level of response during the extinction trials, there are no indications of any actual extinction effects. The Group G *Ss* and the Group C *Ss* more or less maintained the levels attained during the later acquisition blocks. Group S, on the other hand, showed a marked decrement in mean number of past tense verbs between the last acquisition and the first extinction block. The decrement was not merely another block-to-block fluctuation, apparently, as evidenced by the continued low level during the remaining extinction blocks. It is not easy to account for this finding that ceasing to administer shock was accompanied by a greater decrease in the use of past tense verbs than using it as punishment during the acquisition blocks.

Previous research on the relationship between personality characteristics and verbal conditioning behavior has not presented a clear array of results. Taffel (1955), using male psychiatric patients, found that those with high and medium scores on the MAS showed verbal conditioning effects while low MAS scorers did not. On the other hand Daily (1953), using college students, found no difference between high and low MAS scorers in frequency of reinforced response, and Buss and Gerjuoy (1958) found that the mean number of responses increased over the learning trials for the low anxious *S*s, but not for the high anxious *Ss*. As Buss and Gerjuoy indicate, there seems to be no consistent relationship between MAS scores and verbal conditioning.

Cushing (1957) found no relationship between conditionability and a measure of the "need for social approval," defined on the basis of scores on the Edwards Personal Preference Schedule (EPPS). Sarason (1958) evaluated personality characteristics by means of a personality inventory and by ratings in a psychotherapeutic situation by a therapist. He found that high inventory scores on the "anxiety" and "lack of protection" scales were associated with higher verbal conditioning levels, while high "defensiveness" scores went with low conditioning levels. Patients rated as "compliant" by the therapists showed higher conditioning levels than those rated "non-compliant."

Anderson (1958) used scrambled sentences which could be rearranged into either aggressive or neutral statements to test the hypothesis that high response availability leads to more rapid learning in verbal conditioning. He measured aggressive and punishment variables by means of the Thematic Apperception Test, and found significant differences in learning between high and low groups for the following: "punishment for verbal aggression," "total punishment," and "aggression-minus-punishment." In a separate experiment he had *S*s make up sentences when presented with sets of pronouns and verbs to test the hypothesis that *S*s high on "need for approval" (as measured by two EPPS subscores) will condition more rapidly than low *S*s. In accord with the findings of Cushing (1957) this hypothesis was not supported. Kirman (1958) also used a pronoun-verb design and found a positive relationship between "deference need" (EPPS) and verbal conditioning with "good" as reinforcement. He found highly significant negative correlations between rate of learning, using the experimenter response "not-so-good," and the EPPS variables "dominance," "endurance," "heterosexuality," and "aggression."

Before comparing the results of the present investigation with the preceding findings, a word should be said about the differences in manner of statistical analysis. The previous studies have for the most part used some form of spilt-plot design, designated Type I design by Lindquist (1953) and repeated measurements design by Edwards (1950) (or a comparable nonparametric design). In this type of design over-all learning is indicated by a significant column (trials) effect while group differences in learning are shown by a significant row (group) effect, or by a significant trials by groups interaction. Alexander's test is aimed at determining the significance of linear trends and the significance of differences in linear trend. Alexander's test for trend will show insignificance of learning slopes for an infinite array of results which a split-plot type of design will show as significant. The latter is of course sensitive to any difference in means, no matter what the pattern.

Consequently, all comparisons of the present findings with previous research must be considered especially tentative.

Among the many tests in the present experiment of differences in acquisition slope, only those between high and low Mf and Ma scores were significant. The fact that no significant difference was found between *S*s high on Pt and those low on Pt supports the conclusion of Buss and Gerjouy (1958) that there is no consistent relationship between MAS scores and verbal conditioning. This is based upon the very high correlations found between MAS and Pt scores (Brackbill & Little, 1954; Deese, Lazarus, & Keenan, 1953; Eriksen & Davids, 1955). In both the Mf and Ma comparisons, the low scorers showed conditioning while the high scorers showed almost flat curves. The increase for the below median Mf *S*s may have been a function of the female experimenter and their attempts to impress her. Moreover, low scorers on Ma would seem more influenced by environmental cues than those tending toward hypomania, who are characterized by overproductivity in self-initiated thought and action.

On an a priori basis we might expect a positive relationship between verbal conditionability and such characteristics as "suggestibility," "the need for social approval," and "compliance." However, except for the finding of Sarason (1957) that "compliant" patients have higher conditioning levels than "noncompliant" ones, these expected relationships have not been supported. In the present research this negative finding has taken the form of no difference in conditioning rate between high and low Hy scorers.

The EPPS needs which have been found to have significant relationships to verbal conditioning in one or more studies are: "deference," "dominance," "endurance," "heterosexuality," "aggression," and "approval." Among these only aggression, dominance, endurance, and heterosexuality have been shown (Allen, 1957) to have significant correlations with MMPI clinical variables for college males. The reported correlations may be seen in Table 9. Obviously none of the correlations are high enough to place any reliance upon predicting differences in acquisition in verbal conditioning on the basis of MMPI variables.

TABLE 9
REPORTED CORRELATIONS BETWEEN EPPS AND MMPI VARIABLES[a]

EPPS	MMPI	r
aggression	Hy	—.28*
dominance	Si	—.23*
endurance	Hs	—.24*
endurance	Pt	—.24*
endurance	Si	.26*
heterosexuality	D	—.23*

* Significant at .05 level.
[a] As given in Allen (1957).

It is interesting that in the four cases of significant differences in slope during extinction for Group G (Hs, D, Sc and Si) the *S*s above the median showed a tendency to continue increasing in their use of the past tense verbs while the *S*s below the median showed some extinction tendencies (see Figures 4, 5, 6, 7). There is nothing available in the literature (at least to the writers' knowledge) with which to compare these findings of the personality effects upon extinction differences. Moreover, there is nothing apparent in the characteristics of these scales which would account for the differences at extinction, and concomitantly allow for the failure to find differences for these groups during acquisition. Consequently conjecture as to the theoretical significance of these obtained differences will not be attempted.

In the case of the one significant difference in slope for Group S (Hs), the *S*s below the median again decrease in their use of the past tense verbs while the above *S*s increase. But in this case punishment was used during the acquisition trials, so one might have expected the opposite effect from that obtained with "good" as reinforcement. It is not unreasonable to suppose, however, that electric shock is positively reinforcing for the high Hs scorers. That does not account for the decrease for the low Hs scorers, but the decreases for the low Hs scorers in both Groups G and S are quite small and may represent sampling deviations from essentially flat extinction curves.

Finally, not much can be said about the many significant differences in the comparison of means for those above and those below the medians of the various MMPI scales. Since no procedure comparable to covariance analysis is possible with the Alexander test, the significant differences during acquisition may reflect differences in operant level and the significant differences during extinction may reflect acquisition level differences. Consequently, the various operant block means have been presented in Table 6 and the means over the final two acquisition blocks in Table 8 to make evaluation of the obtained significant differences more meaningful. Among the differences during acquisition, only the following are clearly not accountable on the basis of operant level differences: Mf for Group G, and D, Pt, and Si for Group S. In all these cases, there is a shift in the direction of relative magnitude of means from operant block to those over the eight acquisition blocks and the difference between means is larger in the latter case than in the former one. Similarly, the only differences in extinction means which seem clearly attributable to extinction effects are the following: Pa and Pt for Group S. All others show the same direction of differences in means over the last two acquisition blocks and over-all extinction blocks.

E. Summary

The operant conditioning of verbal behavior is frequently considered to have important implications for theories of interpersonal behavior and for psychotherapeutic practice. The present study was aimed at providing information on the effects in verbal conditioning of certain personality variables, a few reinforcement conditions, and their interaction.

All *S*s were males and all were given the MMPI. On each trial in the experiment, *S* was presented with a card containing four pronouns and two different verbs, one in the past tense and one in the present tense. His task was to make up a sentence using one of the pronouns and one of the verbs.

There were 150 trials and these were divided into 15 blocks of 10 trials each.

All *S*s were haphazardly assigned to one of three Groups: G, S, or C. The experimenter said, "good" whenever a Group G *S* made up a sentence using the past tense verb, and administered a mild electric shock whenever an *S* in Group S used the past tense verb. No reinforcement was used with Group C. At the end of the 150 trials, all *S*s in the reinforced groups were asked questions to determine their awareness of the relationship between their responses and reward or punishment. Only those *S*s who did not verbalize this relationship were included in the analyses.

The Group G *S*s did show effects of verbal conditioning during acquisition but the Group S *S*s did not. No extinction effects were found for Group G, and when punishment was stopped for Group S a marked further decrease, rather than increase, in the use of past tense verbs followed. Both Group G and Group S differed significantly from Group C in the mean number of past tense verbs used during the extinction trials.

The median score on each of the 10 most widely used MMPI clinical scales and the median of the Hy-Pt distribution were computed for each group separately, and the learning trends for high (above the median) and low (below the median) groups were compared. Significant differences in linear trend were found in the comparisons involving the Mf and Ma scales for Group G during acquisition, and in the comparisons on Hs, D, Sc, and Si for Group G and on Hs for Group S during extinction. There were a large number of significant differences between group means for the high and low groups on the various scales, but these could not be evaluated unambiguously.

REFERENCES

1. ALEXANDER, H. W. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533-557.
2. ALLEN, R. M. The relationship between the Edwards Personal Preference Schedule variables and the Minnesota Multiphasic Personality Inventory Scales. *J. Appl. Psychol.*, 1957, **41**, 307-311.
3. ANDERSON, D. E. Personality variables and verbal conditioning. Unpublished doctoral dissertation, Univ. of Nebraska, 1958.
4. BINDER, A., MCCONNELL, D., & SJOHOLM, N. A. Verbal conditioning as a function of experimenter characteristics. *J. Abnorm. Soc. Psychol.*, 1957, **55**, 309-314.
5. BRACKBILL, G., & LITTLE, K. B. MMPI correlates of the Taylor Scale of Manifest Anxiety. *J. Consult. Psychol.*, 1954, **18**, 433-436.
6. BUSS, A. H., & GERJUOY, I. R. Verbal conditioning and anxiety. *J. Abnorm. Soc. Psychol.*, 1958, **57**, 249-250.
7. CUSHING, M. C. Affective components of the response class as a factor in verbal conditioning. Unpublished doctoral dissertation, Univ. of Nebraska, 1957.
8. DAILY, J. M. Verbal conditioning without awareness. Unpublished doctoral dissertation, State Univ. of Iowa, 1953.

9. Deese, J., Lazarus, R. S., & Kennan, J. Anxiety, anxiety reduction and stress in learning. *J. Exp. Psychol.*, 1953, **46**, 55-60.

10. Edwards, A. L. *Experimental Design in Psychological Research.* New York: Rinehart & Co., 1950.

11. Eriksen, C. W., & Davids, A. The meaning and clinical validity of the Taylor Anxiety Scale and the hysteria-psychasthenia scales from the MMPI. *J. Abnorm. Soc. Psychol.*, 1955, **50**, 135-137.

12. Greenspoon, J. The reinforcing effect of two spoken sounds on the frequency of two responses. *Amer. J. Psychol.*, 1955, **68**, 409-416.

13. Kirman, W. J. The relationship of learning, with and without awareness, to personality needs. Unpublished doctoral dissertation, Columbia Univ., 1958.

14. Krasner, L. Studies of the conditioning of verbal behavior. *Psychol. Bull.*, 1958, **55**, 148-170.

15. Krasner, L., Weiss, R. L., & Ullmann, L. P. Responsivity to verbal conditioning as a function of two different measures of "awareness." *Amer. Psychol.*, 1959, **14**, 347. (Abstract)

16. Krieckhaus, E. E., & Eriksen, C. W. A study of awareness and its effects on learning and generalization. Paper read at Midwest. Psychol. Ass., St. Louis, May, 1960.

17. Levin, S. M. The effects of awareness on verbal conditioning. *J. Exp. Psychol.*, 1961, **61**, 67-75.

18. Lindquist, E. F. *Design and Analysis of Experiments in Psychology and Education.* Boston: Houghton Mifflin, 1953.

19. Sarason, I. G. Interrelationships among individual difference variables, behavior in psychotherapy, and verbal conditioning. *J. Abnorm. Soc. Psychol.*, 1958, **56**, 339-344.

20. Taffel, C. Anxiety and the conditioning of verbal behavior. *J. Abnorm. Soc. Psychol.*, 1955, **51**, 496-501.

21. Tatz, S. J. Symbolic mediation in "learning without awareness." Paper read at East. Psychol. Ass., Atlantic City, March, 1956.

Department of Psychology
Indiana University
Bloomington, Indiana

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 403-410.

RESPONSE STRENGTH AS A FUNCTION OF RATE OF REWARD*[1]

Department of Psychology, University of South Dakota

Ronald L. Jorgensen and Roger T. Davis

A. Problem

Whenever an investigator varies the amount of a particular external agent, e.g., amount of practice, amount of reward, units of dosage, etc., he unavoidably changes the rate at which the agent is given. In a similar manner the spacing of the administration of the agent covaries with the rate it is given.

The present investigation covaried the number of pellets delivered per bar press and the schedule of reinforcement in a Skinner Box in order to hold rate of reward constant. It asked two questions: (1) if the rate of reward is held constant does a rat respond at a constant rate and (2) is resistance to extinction related to number of pellets that *S* receives or to the number of bar presses.

The above terms are defined in the following manner. Pellets (P) refer to the number of pellets delivered by the mechanism, responses (R) to number of bar presses, and deliveries (D) to the number of times that pellets are delivered. The experimenter specified the number of pellets per delivery (P_d) and the schedule. The latter is the number of deliveries per response (D_r). Reward rate is the product of pellets per delivery (P_d) and deliveries per response (D_r). Response rate (R_t) is the number of responses per unit time. The *E*s predicted in the first question that $P_d \times D_r \times R_t = k$ or $R_t = k/P_d.D_r$.

B. Methods

1. *Subjects*

The *S*s were 24 male albino rats of the Sprague-Dawley strain and were approximately 120 days old when preliminary training was begun.

---

* Received in the Editorial Office on July 12, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This paper represents a revision of a thesis submitted by the first author in partial fulfillment of the requirements for the Master of Arts degree at the State University of South Dakota. The investigation was supported in part by a research grant (M-2258) to the second author from the National Institute of Mental Health, U.S. Public Health Service.

2. *Apparatus*

The *S*s were trained in a Gerbrands Model C Rat Cage with lever. The cage, lever and a Williams-Anger Rotary Pellet Dispenser were mounted in an insulated ice chest having inside dimensions of 20.00 × 13.75 × 11.00 inches. Hereafter this assemblage will be referred to as a Skinner Box. The interior of the Skinner Box was illuminated through a 3.75 inch porthole directly over the rat cage and the porthole was covered with clear plastic.

The delivery of reward to the food cup was controlled by a specially designed counting apparatus which provided the proper number of electrical impulses to operate the pellet dispenser in accordance with the particular program in use. The bar pressing responses of *S*s were automatically recorded on a Harvard Cumulative Recorder and on a direct reading instrument counter. Primary illumination of the room was provided by a single 300 watt incandescent bulb enclosed within the globe of a ceiling fixture. Heavy window shades reduced illumination from daylight. A constant sound screen was provided by the noise of two ventilating fans.

3. *Procedure*

Preliminary training was given over a period of 10 days preceding acquisition, and extinction followed acquisition.

a. Preliminary training. The *S*s were trained during preliminary training to depress the bar in order to activate the pellet delivery mechanism and to raise a hinged plastic lid to obtain a pellet of food in the food cup. A food pellet and honey mixture was smeared on the bar and a programming device automatically delivered a pellet at 2.0 minute intervals during the first three days of preliminary training. During the remaining seven days these aids were not employed.

b. Training during acquisition. Table 1 gives the schedules of reinforcement and amounts of reward per delivery which were assigned to the *S*s during the six conditions of the experiment. Each condition lasted three days and 20 minutes of training was given every day.

The *S*s in Groups A and D received a constant rate of reinforcement, one pellet for every five bar presses. This was achieved by giving *S*s in Group D a 5:1 schedule of reinforcement and one pellet on each delivery under Conditions I through VI and *S*s in Group A a schedule that varied between 5:1 and 80:1 at the same time the number of pellets per reward varied from one to 16. The *S*s in Group C received a constant amount of food per delivery, 16 pellets under all six conditions, and were rewarded at a schedule that expanded between one bar press per delivery under Condition I to 80 bar presses per delivery under Condition VI, and a rate that declined from 16 pellets for five bar presses to 16 pellets for 80 bar presses. The *Ss* in Group B received an expanding schedule between the six conditions, fewer pellets per delivery, and a very sharply attenuated rate—from 16 pellets for five bar presses to one pellet for 80 bar presses.

TABLE 1
SCHEDULES OF REWARD

	Conditions						
	I	II	III	IV	V	VI	VII
Daily trials	1-3	4-6	7-9	10-12	13-15	16-18	19-21
Bar press delivery ratio	5:1	10:1	10:1	20:1	40:1	80:1	extinction
Group A	1	2	4	4	8	16	0
Group B	16	8	4	4	2	1	0
Group C	16	16	16	16	16	16	0
	5:1	5:1	5:1	5:1	5:1	5:1	extinction
Group D	1	1	1	1	1	1	0

The values within the cells represent the number of 0.045 gm. food pellets given in each delivery.

Exceptions to the foregoing conditions occurred during Condition III. This condition was not planned when the experiment was designed but occurred because of the malfunctioning of the programming equipment. In the original design the schedule of reinforcement planned for the third condition was 20:1 for Groups A and B instead of 10:1.

c. Extinction trials. All groups of *Ss* were given extinction trials under conditions identical to the last day of the preliminary training trials with the exception that no food was delivered at any time. The pellet dispenser operated in the same manner as during preliminary training and the noise of the dispenser as well as that of other components of the apparatus was available as a secondary reinforcing agent.

d. Feeding. Each *S* was given a complementary feeding of standard laboratory ration twenty minutes after being returned to the living cage. The amount of this feeding was determined by the amount of food received in the experimental situation and was controlled to maintain the rat at a constant weight level. Water was given ad libidum in the living cage but was not available during experimental sessions.

C. RESULTS

The Sign test was employed to assay the significance of differences between repeated measures on *S*s in a particular group. A difference was significant at the 5.0 per cent confidence level if all six *S*s in a group showed a change in performance in the same direction from one condition to another. Differences between the performance of *S*s in two different groups were tested for significance using the Mann-Whitney *U*-test.

1. *Acquisition*

The average number of bar pressing responses made on each of the 18 days of acquisition are shown in Figure 1 and the average number of pellets received by *S*s in Figure 2. If the prediction that the total number of pellets delivered to an *S* during each session is a constant, all values in Figure 2 should be the same. Within extreme limits of nonreward *S*s manage to receive a nearly constant number of pellets. Inspection of Figure 2 indicates that *S*s in the different groups approach a similar maximum limit for the consumption of pellets during one 20-minute period, but reach this limit during different periods of acquisition. Animals in Group B reach this maximum, on the average, after 2.5 days of acquisition, Group C—6.3 days, Group A—10.5 days and Group D—15.7 days. Of the six possible com-

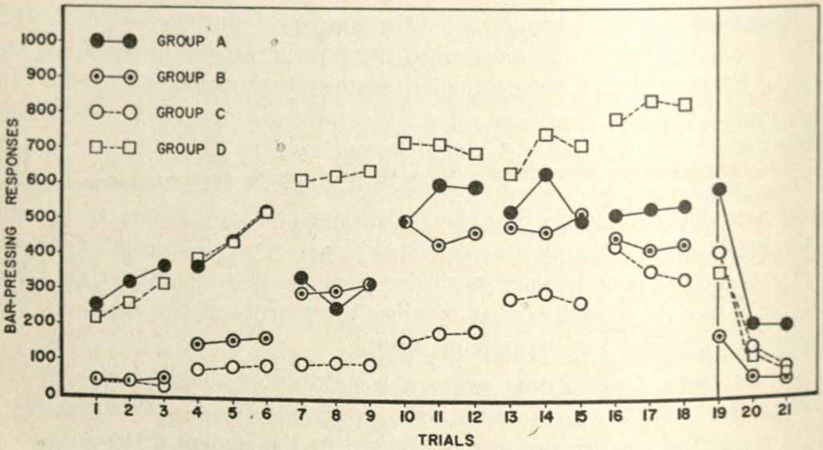


FIGURE 1

Average number of responses made during each daily trial by *S*s of each of the four groups. Breaks in the functions occur between each of the six conditions comprising the period of acquisition and the vertical line designates the beginning of extinction training.

parisons between the number of days that *S*s in the four groups reached maximum reward, only the one between Groups A and C provided a difference that failed to be significant.

The average of the highest number of pellets *S* received during any one of the 18 days was 169.0 gm. and none of the groups differed significantly in this measure. Since each pellet weighs .045 gm., the average maximum intake of dry food was 7.6 gm. An adult rat eats 10-15 gm. food per day and gains weight. During the periods *S*s were near their maximum intake very small amounts of supplemental feeding were necessary to keep each *S*'s weight constant.

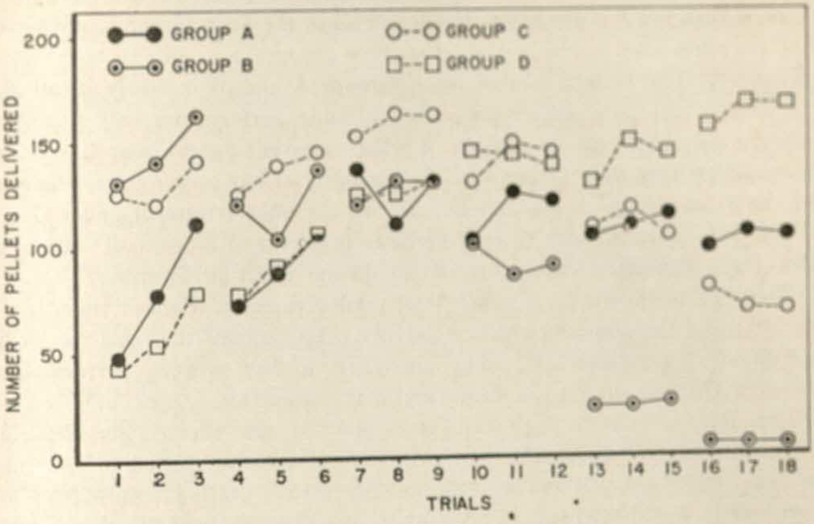


FIGURE 2
Average number of pellets received during each daily trial by *S*s in each of the four groups. Breaks in the functions occur between each of the six conditions comprising the period of acquisition.

During the first three days of the experiment *S*s in Groups A and D press the bar comparably often and *S*s in both groups press the bar significantly more often than *S*s in either Group B or C. However, during this period *S*s in the latter two groups receive significantly more pellets than *S*s in either Group A or D.

Between Condition I and II *S*s in all four groups increase bar pressing and Ss in Group B press the bar significantly more often than *S*s in Group C. The latter two groups both experienced a change in schedule from 5:1

to 10:1, but *S*s in Group B received half as many pellets per bar press as they had received in Condition I. The performance of *S*s in Groups A and D were nearly identical and did not overlap with the performance of *S*s in the other two groups. The only significant difference in pellets received during Condition II was between *S*s in Groups A and C.

Under Condition III all of the Ss in Group D press the bar more frequently than any *S* in the other groups and all of the *S*s in Group C press the bar less often than any *S* in the other three groups ($p < .001$ in both comparisons). The performance of *S*s in Groups A and B is nearly identical. In spite of the gross differences in bar pressing performance only one significant difference exists in number of pellets received; *S*s in Group C receive more pellets than *S*s in Group D. The similarity of bar pressing performance of the *S*s in Groups A and B which receive the same treatment under Condition III, indicated that prior differences in rate and amount of food per delivery either cancel each other out or do not effect performance.

During Condition IV the schedule of reinforcement is changed from 10:1 to 20:1 and the number of pellets per bar press remains the same for each group as in Condition III. The differences in bar pressing performance between the four groups are significant in the same instances as under Condition III, but rats in both Groups A and C press the bar significantly more often in Condition IV than III. This is the first period of training where *S*s in Group B did not increase their bar pressing significantly to compensate for the reduction in number of pellets received per bar press. During Condition IV *S*s in Groups C and D received a comparable number of pellets, those in Group A significantly fewer than those in D, and *S*s in B significantly less than *S*s in either C or D. None of the groups received significantly more pellets in this period than in a prior period but *S*s in Group B received significantly fewer than they had received in the first period of training.

During Condition V, *S*s in Groups A, B, and C were on a 40:1 schedule of reinforcement. The *S*s in Group C press the bar significantly more often than in any prior period of training, but the increase is not sufficient to keep up with the number of pellets delivered to *S*s in Group D. The only significant differences in number of pellets received are those between the *S*s in Group B and *S*s in all of the other groups.

In the final Condition, VI, two groups, A and C, are on the same schedule of reward and receive the same number of pellets per delivery. The *S*s in three groups, A, C, and D, all are rewarded at the same rate. The *S*s in Group D press the bar significantly more often than those in any of the other groups, and the *S*s in Group A press the bar significantly more than those in C. All groups differ significantly from one another in number of pellets received.

2. *Extinction*

On the first day of extinction, only the *S*s in Group D press the bar significantly less than during the last day of acquisition, and throughout the first three days of extinction the *S*s in Group A showed significantly greater resistance to extinction than *S*s in any of the remaining groups.

D. DISCUSSION

The similarity from *S* to *S* of the maximum number of pellets obtained on a single day of acquisition suggests that each *S* was operating under the same drive. This drive resulted from 24 hours of food deprivation and maintenance of constant body weight. The *S* is given only 20 minutes to procure the amount of food indicated by the drive and any factor which interferes with the efficiency of *S*'s response reduces the likelihood that the optimum amount will be reached.

One factor that was present in all conditions and that undoubtedly limited responsiveness in rats was the unavailability of water during the training sessions. Rate of reward restricts the number of pellets obtained and restriction depends on the interaction of the two components of rate of reward, schedule and number of pellets per delivery. Changes in the schedule of reinforcement alters the spacing between rewards and when spacing reaches the point that *S* cannot obtain the amount dictated by the drive, extinction begins. Extinction is evident in the performance of *S*s in Group B during Conditions V and VI and is further illustrated by the minimal responsiveness by *S*s in Group B during the extinction period. The number of pellets per delivery determines the duration of the period of food consumption. When *S* presses the bar five times and receives one pellet, the pellet is ingested immediately and *S* resumes bar pressing. During the next five responses the pellet is eaten and swallowed. The next pellet is then delivered and consumed immediately. If *S* receives a large number of pellets in a delivery, it must stop and consume the pellets between periods of bar pressing. Pausing reduces the time available to press the bar. The number of pellets per delivery is not a limiting condition unless the schedule involves considerable spacing between deliveries. This is shown in Figure 2. The *S*s in Group A obtain fewer pellets under Conditions IV, V, and VI than *S*s in Group D. However, *S*s in Groups B and C obtain the maximum number of pellets dictated by 24 hour drive during Condition I in which 16 pellets are given for five bar presses.

This mechanical account of performance is adequate to describe acquisition of bar pressing, but inadequate to account for *S*s performance during the period of extinction. During this period the *S*s in Group A show significantly greater resistance to extinction than *S*s in either of the remaining groups. Since Groups A and D were rewarded at the same rate throughout the experiment the difference cannot be ascribed to the effects of different rate of reward. The *S*s in Groups A and C received the same number of pellets per delivery during Condition VI. In all the conditions prior to Condition VI *S*s in Group C received more pellets per delivery than *S*s in Group A. Since *S*s in both of these groups received the same schedules of reinforcement the difference between these groups cannot be related to different rates of reward or either of the components of rate. One would expect that *S*s in Group C rather than Group A should show the greatest resistance to extinction on the basis of the total pellets received by *S*s in these two groups.

The superiority of the performance of *S*s in Group A during the extinction period cannot be related to expectancy of reward since *S*s in Group C received more reward per delivery in five of the six conditions during acquisition. The only difference between the treatment of Groups A and C that might dispose *S*s in Group A to more resistance to extinction than *S*s in Group C is that *S*s in Group A received an increasing number of pellets per delivery from condition to condition.

E. Summary

Four groups of six rats each were given experimental training in a Skinner box under different rates of reward. The rate was determined by the amount of reward in a single delivery and the ratio of bar presses to delivery of reward.

1. The *S*s tended to approach an optimum level of reward for an experimental period. Any change in condition brought about an immediate response in bar pressing which tended to supply this level of reward.

2. Departures from the optimum reward level largely were due to extreme variations in rate of reward.

3. Resistance to extinction appears to depend on prior experience with amount of reward per delivery.

Department of Psychology
Purdue University
Lafayette, Indiana

Department of Psychology
University of South Dakota
Vermillion, South Dakota

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 411-422.

STUDIES ON TRANSFER OF BRIGHTNESS DISCRIMINATION FROM BAR-PRESSING TO A Y MAZE*[1]

Department of Psychology, New York University

M. R. D'AMATO AND H. JAGODA

A. INTRODUCTION

In two previous studies (2, 3) it was shown that extensive early (post-weaning) training on a brightness discrimination in a bar-pressing situation later produced specific transfer effects on brightness discrimination acquisition in a Y maze, whether testing on the Y maze occurred soon after acquisition of the initial discrimination (3) or was delayed until early adulthood (2). The *S*s that had a bright light as S^D in the original discrimination (dim light as S^{Δ}) later learned the Y maze discrimination problem (bright positive) significantly faster than *S*s for which the dim light served as S^D in the initial discrimination. Control *S*s that received no prior discrimination training at all learned the Y maze discrimination at an intermediate rate. It was concluded that the extended discrimination training in the Skinner boxes served to develop specific response tendencies toward the discriminanda (e.g., approach and avoidance), and that these response tendencies mediated the subsequent transfer effects. Had the establishment of such general perceptual factors as "paying attention to visual cues" been of primary importance, both experimental groups, it was argued, would have been superior to the control group on the Y maze discrimination.

B. EXPERIMENT 1

1. *Problem*

On the assumption that the observed transfer from the Skinner box to the Y maze was mediated by specific associations to the discriminanda, for example, approaching bright light and avoiding or withdrawing from dim illumination, any operation which conditions these associations equally to both discriminanda should prevent transfer between the Skinner box and Y maze discriminations. One such operation consists of alternating the roles of the bright and dim illumination, i.e., repeated reversal training. In Experiment 1, therefore, the experimental *S*s were subjected to repeated reversals of a brightness discrimination in the Skinner boxes. They were later trained on the Y maze on a brightness discrimination with bright positive and reversed after achieving criterion. On the supposition that the extensive early reversal learning would preclude the establishment of differential specific associations to the discriminanda, and assuming further that the Skinner box discrimination training does little in the way of developing general perceptual factors, the experimental *S*s should learn the Y maze discrimination no faster than control *S*s that received no prior discrimination experience. It is not clear how the early Skinner box reversal training should affect reversal learning on the Y maze, though on the basis of the specificity of the transfer obtained on discrimination acquisition in the Y maze, it is improbable that the early experience with reversal learning would have a facilitating effect on reversal learning in the Y maze.

---

* Received in the Editorial Office on July 13, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was supported by Grant M-2051 from the National Institute of Mental Health, National Institutes of Health, United States Public Health Service, and Grant G-14724 from the National Science Foundation.

2. *Method*

a. Subjects and apparatus. The *S*s were 24 pigmented rats (hooded and complete blacks and browns; 15 males and 9 females), and 24 albinos (9 males and 15 females) taken from litters bred in our laboratory. The pigmented and albino *S*s were run as separate replications.

The Skinner boxes consisted of a rack of modified rat cages, with a bar protruding into the rear of the cage and a dipper arrangement for delivering water reward; a 1-inch pilot light was centered above the bar and dipper opening. Each modified cage alternated with an unmodified (control) cage, and while no attempt was made to acoustically insulate the cages, precautions were taken to minimize the possibility that a discrimination would be formed on the basis of such cues as dipper noises. (Further details are given in a previous report (3).)

Transfer testing took place on four automatic Y maze discrimination apparatuses described in a previous report (4).

b. Skinner box discrimination training. Within 48 hours after weaning, which took place at 21 days of age, pups were removed from litters and, balancing for sex, randomly assigned between Groups C(P) and E(P) or Groups C(A) and E(A), (P) and (A) referring to the pigmented and albino replications. After assignment to groups, the control *S*s (Groups C(P) and C(A)) and experimental *S*s (Groups E(P) and E(A)) were individually placed in unmodified and experimental cages, respectively, where they remained for 26 days. During this time the control *S*s were fed and watered ad lib.

The discrimination procedure for the experimental groups was as follows. On Days 1 and 2, S^D was continuously present and all bar-presses resulted in water delivery; water was not otherwise available. On Days 3 and 4, a 5-minute S^D period alternated regularly with a 5 minute S^Δ period, a continuous reinforcement schedule (crf) prevailing in S^D. During Days 5 to 26, two nonoverlapping 3-minute S^D periods occurred randomly in every hour period, the remaining 54 minutes of the hour being spent in S^Δ; the schedule during S^D was always crf. The discriminanda consisted of two brightness levels of the pilot light (the same as used in the previous studies and specified in (3)). The program was in effect for 20 hours of each day, and the remaining 4 hours were spent in darkness, with the dippers and response relays disconnected.

The present procedure was similar in all respects to that used in the D'Amato (2) study, except that the bright and dimmed pilot light alternated daily as S^D. Hence, the *S*s of this study were subjected to repeated reversal learning.

On Day 26, *S*s were removed to general laboratory cages, with food and water constantly available. About five weeks later, when *S*s were 85 days of age, water was removed and our standard 4-day 23-hour water deprivation regimen begun.

c. Y maze discrimination and reversal training. As in the previous studies, all *S*s were trained, 20 trials per day, on a simple brightness discrimination problem with bright (10.0 ft.-c.) positive and dim (0.01 ft.-c.) negative, until the criterion of 18 correct responses out of any 20 successive trials was met.

On a correct choice the sequence of events was as follows. The *S* was permitted 1.75 seconds ± .25 seconds drinking, after which the dipper retracted out of sight. The discriminandum remained on in the arm for an additional 8 seconds. Then a 57 second ± 3 second intertrial interval followed, during which all lights in the maze were extinguished. Finally, a flashing light during a 4.5 second ± .5 second interval signalled the beginning of the next trial. On an incorrect choice the sequence of events was the same, except that the dipper was retracted during the initial 1.75 second interval. Fifteen to 30 minutes after a day's trials, *S* was given access to water for one hour. The estimated mean deprivation duration at the start of a day's trials was 22½ hours.

The day after achieving acquisition criterion *S*s were placed on reversal learning, and except for the reversal of the positive and negative stimuli, the procedure employed was the same as that of acquisition.

3. *Results*

a. Skinner box discrimination. In order to evaluate the discriminatory behavior of the experimental *S*s during the first phase of the experiment, a discrimination ratio (DR) was calculated for every *S* on each day. The DR was defined as the ratio of the rate of responding in S^{Δ} over the rate of responding in S^{D}. The mean DR over the last 10 days of discrimination training provided a single index of each *S*'s discriminatory behavior. Group means based on this measure were quite similar for Groups E(P) and E(A), viz., .087 and .082, respectively, indicating very little difference between the groups in the terminal level of discriminatory performance. These group means, it is interesting to note, are close to the values (.050 and .077) obtained in the D'Amato (2) study, indicating that the experimental *S*s were discriminating adequately by the end of discrimination training.

b. Y maze discrimination acquisition and reversal. The mean numbers of trials to acquisition criterion for Groups C(P), E(P), C(A), and E(A)($N = 12$ per cell) were 54.9, 68.1, 102.7, and 99.4, respectively. A 2×2 analysis of variance showed the strain variable alone to be significant ($F = 26.4$, $P < .001$). The pigmented *S*s, as is apparent from the group means, learned the Y maze discrimination considerably faster than the albinos, a result previously reported (5). Far from learning the Y maze discrimination faster than the controls, the experimental *S*s were actually somewhat inferior to the latter, 78.8 and 83.8 being the respective group means.

In order to determine whether goodness of discrimination was related for the Skinner box and Y maze discriminations, a rank order correlation was calculated between the DR scores reported above and the mean numbers of trials to acquisition of the Y maze discrimination. This was done separately for the pigmented and albino *S*s, and yielded the values of —.216 and —.265, both values falling far short of significance.

The mean numbers of trials to reversal criterion for Groups C(P), E(P), C(A), and E(A) were, in order, 141.1, 137.6, 163.7, and 146.5. All *F*s of a 2×2 analysis of variance fell far short of significance. The early Skinner box experience with repeated discrimination reversal, therefore, did not facilitate reversal learning on the Y maze.

4. *Discussion*

The failure of the experimental *S*s to show transfer effects on acquisition of the Y maze discrimination supports the hypothesis that the transfer obtained in previous studies (2, 3) was mediated by specific response tendencies and that the development of general perceptual sensitivity does not occur in the present situation, even after prolonged and extensive early discriminatory experience. The comparability of the DRs obtained in this study with those reported in the earlier work rules out the possibility that *S*s of this study did not show transfer effects because they failed to develop adequate discriminations in the Skinner boxes. Further, since larger sample sizes were used in the present study than in the earlier work, the probability of detecting a transfer effect, if it existed, was in this experiment at least as good as in the previous studies.

The question arises as to the precise nature of the specific response tendencies or associations that mediate transfer from the Skinner box discrimination to the Y maze. One possibility was suggested in a previous paper (2). Prolonged training in the Skinner box with the bright light as S^D might ultimately result in *S*'s being conditioned to approach the bar and dipper (reward) area of the cage on the stimulus change dim-bright (i.e., S^Δ—S^D). When later tested on the Y maze with bright positive, *S* would be exposed to the same temporal pattern dim-bright when he looked into the negative and positive arms in sequence. Because of the similarity in stimulus patterns, the approach response to the reward area of the bright (positive) arm should transfer to some degree, thereby facilitating discrimination learning. Avoidance or withdrawal responses could become conditioned to the bright-dim pattern in a similar fashion, i.e., *S* would eventually become conditioned on the stimulus change bright-dim to withdraw from the bar and dipper area to other parts of the cage. When this stimulus pattern (generalized) is again encountered in the Y maze, *S* could be expected to withdraw from the dim region (arm), increasing his chances of finally choosing the bright (positive) arm. Both of these specific response tendencies would serve to facilitate discrimination learning on the Y maze, since the brightly illuminated arm was positive. A similar analysis would, of course, account for negative transfer on the Y maze discrimination in the case of *S*s having the dimmed pilot light as S^D in the Skinner box discrimination training.

One may develop the preceding interpretation into perhaps even greater specificity. Considering *S*s having the bright light as S^D, in approaching the bar and dipper area on the stimulus change dim-bright these *S*s quite literally approached the bright light, since the stimulus source, the pilot light, was situated directly above the bar and water dipper. Positive transfer on the Y maze may then be understood largely in terms of conditioned approach to the bright light source (rather than conditioned approach to a potential reward area), and negative transfer as conditioned approach to the dim source. If this were the case, changing the location of the pilot light from the rear to the front of the experimental cages should have the effect of reversing the direction of the transfer obtained on the Y maze, since with the pilot light on the front wall of the Skinner boxes *S*s would have to withdraw from S^D in order to get to the bar and dipper area. In Experiment 2 this possibility was explored.

C. EXPERIMENT 2

1. *Method*

a. Subjects and apparatus. The *S*s were 20 albino rats (11 males and 9 females) from two litters of 10 pups each bred in our laboratory.

Except for the removal of the pilot light from the rear of the experimental cages to the front wall, all apparatus was the same as used in the previous study. It should be specified that since the pilot lights were powered through the same 28 VDC power source that supplied the dippers and programming relays, the discriminanda flickered any time an *S* made a bar-press, a condition that prevailed in all preceding studies as well.

b. Skinner box and Y maze discrimination training. Four of the pups (21 to 23 days of age) of each litter were randomly assigned to Group C (controls) and individually placed in the unmodified cages. Three pups each were assigned to Groups B and D and placed in experimental cages. The *S*s of Group B had the full-bright pilot light as S^D and dimmed pilot light as S^Δ, the reverse holding true for *S*s of Group D. The Skinner box discrimination procedure was the same as employed in the D'Amato (2) study and differed from that of Experiment 1 only in that there was no reversal training.

The Y maze training, which started when *S*s were 90 days of age, followed closely the procedure employed in Experiment 1, with the single change that the acquisition criterion was made more stringent by requiring that the last 10 criterional trials be correct.

2. *Results*

a. Skinner box discrimination. As in Experiment 1, DRs were calculated for all experimental *S*s on each of the last 10 days of Skinner box training and the mean DR for the 10 days obtained for each *S*. The group means based on these measures were .035 and .092 for Groups B and D, respectively. Despite the small sample sizes (six *S*s per group), the difference between the groups proved significant by the U test ($P < .05$). In the earlier study (2) to which this experiment is most comparable, Group B had a significantly *larger* mean DR than Group D (.077 vs. .050). It was suggested there that the difference between the groups in mean DRs was due to *S*'s tendency to avoid the bright light by retreating to the opposite wall of the cage. Since in the latter study the pilot light was located in its usual place above the bar, responding during S^D was thereby impeded in *S*s of Group B, whereas responding during S^Δ was discouraged in *S*s of Group D. The results of this experiment verify this explanation, inasmuch as the changed location of the pilot light to the front wall produced the predicted reversal in the difference between the mean DRs of Groups B and D.

b. Y maze discrimination. The mean numbers of trials to criterion for Group B ($N = 6$), Group C ($N = 8$), and Group D ($N = 6$) were, in order, 79.8, 102.3, and 100.7, none of the differences achieving significance. While statistical significance could most probably be achieved by increasing the sample size through replication, the direction of the differences made this line of attack seem unprofitable. It seemed clear from the results that changing the location of the discriminanda source does not reverse the direction of the transfer effects on the Y maze.

(It should be mentioned that the results of the present and of other, unpublished, work suggest that less transfer occurs with the discriminanda source on the front wall, i.e., away from the dipper and bar region of the cage. Another variable which appeared to have an effect on the extent of transfer to the Y maze discrimination was the presence or absence of flicker in the discriminanda source of the Skinner boxes. As mentioned above the present as well as the earlier published studies were conducted with the pilot lights powered through the 28 VDC source, producing flicker in all pilot lights with every bar-press. In intervening studies the flicker was removed by having the lights powered by a separate (AC) source. It was our impression, bolstered by some data, that transfer was enhanced when flicker occurred, possibly because such flicker increased attending to the stimuli in the Skinner boxes. In any event, there seemed little doubt that transfer was increased by permitting flicker in the Skinner box stimuli and particularly by locating the discriminanda source close to the bar and dipper area.)

3. *Discussion*

While the results of the present study do not support the view that transfer of brightness discrimination from the Skinner box to the Y maze is based on the conditioning of approach responses to S^D, the possibility still exists that the conditioning of approach (and withdrawal) responses to the bar and dipper area, as previously described, is the important factor. Since, presumably, these responses would be conditioned to the appropriate dim-bright or bright-dim stimulus pattern, they would be relatively independent of such variables as location of the discriminanda source, though no doubt the extent of the transfer obtained, but not the direction, might be influenced by the positioning of the stimulus source. Other kinds of specific associations to the discriminanda are no doubt possible, e.g., "anxiety" and "frustration" responses, and could possibly mediate the transfer effects.

D. Experiment 3

1. *Problem*

It is conceivable that prolonged early discrimination training results primarily in the establishment of specific response tendencies toward the discriminanda, rather than in the development of heightened perceptual sensitivity of a general nature, only if the discrimination training occurs under quite constant conditions, as was the case, for example, in Experiment 2 and the previously reported studies (2, 3). Variations in the properties of the relevant stimuli could possibly have the effect of increasing the latter at the expense of the former. And with increased perceptual sensitivity (e.g., increased ability to seek out and attend to relevant cues) one might reasonably expect an enhancement of the amount of transfer shown on the Y maze discrimination. As an initial test of this latter hypothesis a study was initiated in which the brightness of S^D was varied periodically in one group of *S*s, while for a second group S^D was held constant at a fixed brightness level. The point of interest was whether the varied S^D subjects would show more (positive) transfer effects than the constant S^D animals.

2. *Method*

a. Subjects and apparatus. The *S*s were 44 albino rats (22 males and 22 females) obtained from five litters bred in our laboratory.

The apparatus was the same as used in the previous studies, with the pilot lights of the Skinner boxes being returned to the rear wall.

b. Skinner box and Y maze discrimination training. Within 1 to 7 days after weaning, littermates were divided among Groups BC, BV, and C and placed in the experimental or control (unmodified) cages. Group C constituted the control group. Group BC received the usual discrimination training with the standard (full) bright pilot light as S^D; however, for S^Δ the pilot light was extinguished, rather than dimmed as in previous studies. For Group BC, then, the Skinner box discrimination training occurred with S^D (and S^Δ) held constant. A set of three brightness values served as

S^D for Group BV. The brightest (B1) was the full-bright pilot light; the dimmest (B3), obtained by placing an 820 ohm resistor in series with each pilot light, was the value that in former studies served as S^Δ; a third, intermediate brightness (B2) was obtained by the use of a 470 ohm dropping resistor. The extinguished pilot light again acted as S^Δ. Thus, Group BV acquired the Skinner box discrimination with an S^D that varied over a set of three values and a constant S^Δ.

The Skinner box discrimination training generally followed the procedure of Experiments 1 and 2. On Days 1 and 3 Group BV was exposed to brightness B1, to B2 on Day 2, and to B3 on Day 4; on Days 5 to 26 the three brightness values were assigned in blocks of three, each brightness represented within each block.

Training on the Y maze started when *Ss* were 90 to 95 days of age, the discrimination procedure being identical with that of Experiment 1.

3. *Results*

a. Skinner box discrimination. The usual DRs were calculated for *Ss* of Groups BC and BV ($N = 12$ per group). The mean DRs for the last 10 days of discrimination training were .073 and .085, respectively, the difference between the group means being small and statistically insignificant.

b. Y maze discrimination. The mean numbers of trials to acquisition criterion for Groups BC, BV, and C ($N = 20$) were 65.3, 77.8, and 94.2, respectively. An analysis of variance applied to the data produced an F (4.52) for conditions significant at the .01 level. The difference between the means of Groups BC and C was quite significant ($t = 2.94$, $P < .01$), but the Group BV vs. Group C difference fell short of significance ($t = 1.67$, $P \approx .10$). It seems clear, therefore, that the presentation of a varied S^D during the initial discrimination training does not enhance transfer to the Y maze discrimination; if anything, it leads to a reduction in the amount of positive transfer obtained. Whether such conditions facilitate the development of general perceptual factors is still an open issue.

E. GENERAL DISCUSSION

It is useful to distinguish among studies of transfer in discrimination learning in accordance with whether the transfer involves homogeneous or heterogeneous situations. Probably most relevant researches have employed homogeneous situations, i.e., the instrumental response and experimental situation remain relatively constant, but the discriminanda vary from problem to problem (e.g., Harlow's work with "learning sets"). In heterogeneous situations, on the other hand, the initial and transfer discriminations involve different instrumental responses and, quite likely, different experimental situations as, for example, transferring from a discrimination box to a jumping stand. In homogeneous situations, with the instrumental response and experimental situation held constant, nontrivial cases of transfer can be studied only by varying the discriminanda, while with heterogeneous situations studies of discrimination transfer may or may not involve changes in the relevant stimuli.

Obviously, any degree of heterogeneity can exist between the initial and transfer discriminations. The series of studies of which the present experiments form a part involved considerable heterogeneity, since both the instrumental response and experimental situation differed rather widely in the initial and transfer discriminations. The Skinner box discrimination required a bar-pressing response and successive discrimination; the Y maze discrimination involved the simple response of progression and required a choice between alternatives, i.e., simultaneous discrimination. That significant transfer effects with rats can be obtained in the face of such heterogeneity (even after considerable delay between the initial and transfer discriminations) has not often been demonstrated previously and is of some interest in and of itself.

Apart from the latter, the major finding of our work in this area is that the transfer effects obtained were quite specific to the relationship holding between the discriminanda of the initial and transfer discriminations. Significant development of general perceptual sensitivity did not occur even though the initial discriminatory experience (*a*) was prolonged and intensive, (*b*) occurred early in the organism's life, while testing was delayed to early adulthood (2), and (*c*) involved the varied experience with the relevant stimuli provided by repeated reversal learning (Experiment 1).

The basis of the specific transfer effects which we have observed in our research is still difficult to specify. Having ruled out in Experiment 2 conditioned approach responses to S^D as an important factor, the possibility yet remains that approach (and avoidance) associations are formed to the bar and reward area in the manner described above. In our view this interpretation seems the most promising one now available.

Certain other possibilities are easily ruled out. It is not difficult, for example, to demonstrate that mere exposure to the stimuli during Skinner box training cannot account for the observed transfer effects, at least not through the mechanism of altering the rat's normal negative phototropism. Analysis of the first five discrimination trials on the Y maze reveals that there is little difference between the per cent correct responses (i.e., responses to the brighter arm) made by the control and experimental *S*s. The 28 control *S*s of Experiments 2 and 3 above and the 14 controls of a previous study (2) combined averaged 44 per cent correct responses on the first five Y maze trials. The corresponding combined figure for the 20 Group B *S*s of the latter study and of Experiment 2 was 42 per cent; the 20 Group D *S*s of these same two studies also averaged 42 per cent correct responses. There is, therefore, no evidence for the operation of differential negative phototropism on the early Y maze discrimination trials. Further, it has been shown (2) that control *S*s that work for water in the Skinner boxes under the same stimulus programs as experimental *S*s (but without discrimination training) do not later show transfer effects on the Y maze. It is clear, in conclusion, that the prior discrimination training is necessary for the appearance of transfer effects on the Y maze.

Finally, the absence of significant development of perceptual sensitivity of a general nature in spite of the prolonged and intensive early discrimination training has implications for the important issue of the role of early experience in shaping adult behavior organization. While it seems well established that severe early deprivation of basic visual experiences can result in impaired sensory and intellectual functioning in adulthood, it is not clear that, given a certain basic minimum of such experiences, further "enrichment" will lead to significant improvement in adulthood of such intellectual functions in the rat as discrimination learning, maze learning, and problem solving. Earlier studies (e.g., 1, 6, 7) which seemed to indicate that rats raised in a "rich" environment displayed as adults more maze and problem-solving ability than *S*s raised in a restricted environment have recently been reinterpreted. It now appears (e.g., 8) that the differential early experience affected directly the rat's exploratory drive, rather than its maze or problem-solving ability, and that the inferior performance of the "deprived" *S*s was simply the result of a heightened, competing, exploratory drive.

Our results have a similar significance. Extensive early discrimination training has been shown to affect adult discrimination learning in an entirely different situation, not by increasing the discrimination-learning ability of the rat, but rather by establishing specific associations capable of transferring to the later discrimination problem.

F. SUMMARY

Three experiments were conducted on the effect of prolonged and intensive early Skinner box brightness discrimination training on the acquisition in early adulthood of a brightness discrimination in a Y maze. In Experiment

1 repeated brightness reversal training in the Skinner box was shown to have no effect on the later acquisition and reversal of a brightness discrimination on the Y maze, supporting the view that the transfer obtained in earlier reported studies was based on specific response tendencies (or associations), rather than on general perceptual factors. In Experiment 2 these specific response tendencies were shown not to depend on the conditioning of approach responses to the discriminanda *per se*. The results of Experiment 3 indicated that more positive transfer could be obtained by using constant training conditions (i.e., constant S^D) during the early Skinner box discrimination learning experience than by systematically varying S^D over a set of three brightness values. The implications of the experimental results for the problem of the role of early experience in determining adult behavior organization was discussed.

REFERENCES

1. BINGHAM, W. E., & GRIFFITHS, W. J. The effect of differential environments during infancy on adult behavior in the rat. *J. Comp. Physiol. Psychol.,* 1952, **45**, 307-312.
2. D'AMATO, M. R. Transfer of an early acquired brightness discrimination from bar-pressing to a Y maze. *J. Comp. Physiol. Psychol.* (in press).
3. D'AMATO, M. R., & JAGODA, H. Transfer of brightness discrimination from bar-pressing to a Y maze. *Psychol. Rep.,* 1960, **7**, 83-90.
4. ———. Effects of extinction trials on discrimination reversal. *J. Exp. Psychol.,* 1960, **59**, 254-260.
5. ———. Effect of early exposure to photic stimulation on brightness discrimination and exploratory behavior. *J. Genet. Psychol.,* in press.
6. FORGAYS, D. G., & FORGAYS, J. W. The nature of the effects of free-environmental experience on the rat. *J. Comp. Physiol. Psychol.,* 1952, **45**, 322-328.
7. HYMOVITCH, B. The effects of experimental variations on problem-solving in the rat. *J. Comp. Physiol. Psychol.,* 1952, **45**, 313-321.
8. WOODS, P. J., FISKE, A. S., & RUCKELSHAUS, S. I. The effects of drives conflicting with exploration on the problem-solving behavior of rats reared in free and restricted environments. *J. Comp. Physiol. Psychol.,* 1961, **54**, 167-169.

Department of Psychology
New York University
NewYork 53, New York

Published as a separate and in *The Journal of Psychology*, 1961, 52, 423-427.

VOLUME-UNIT GRAPHS: THEIR PRODUCTION AND APPLICABILITY IN STUDIES OF INFANTS' CRIES*[1]

Long Island Jewish Hospital and Hunter College

Vincent R. Fisichelli, Samuel Karelitz, John Eichbauer
and Laura S. Rosenfeld

Studies of infant vocalizations currently in progress at the Long Island Jewish Hospital (1) are focused upon a variety of hypotheses, one of which is concerned with differences in crying behavior between normal infants and those suffering from brain disorders. Observations by one of the present authors (S. K.), dating back more than thirty years, suggested that infants with brain damage cry less frequently and less actively than normal children do, and that their cries are audibly different from those of normal infants. Data bearing on this problem have been collected during the past four years and, as of this writing, more than 1300 samples of crying have been taken from normal, abnormal and prematurely born infants. The cries were magnetically tape-recorded on an Ampex 601. The microphone used was an electrovoice 666. These data are being analyzed in several ways. The purpose of this paper is to describe a technique similar to ones already described elsewhere (2, 3), for converting the tape recordings into printed records through a graphic milliameter and to indicate how these graphs supply values for comparing the vocalizations of normal and brain involved infants.

The Ampex 601 magnetic tape recorder-player is equipped with a volume-unit meter which permits visual monitoring of the signal. The deflections of the needle in this meter can be recorded graphically if the machine is put in circuit with a graphic-milliameter (Esterline-Angus, Model AW). In the present instance the signal was amplified before being relayed to the graphic-ammeter. This provided a variable control of the signal entering the ammeter. The Ampex 620 speaker served admirably for this purpose. Since the Esterline-Angus model used in this study is a D.C. meter (A.C. models are available), a meter rectifier was inserted between it and the amplifier. The final *V-U* graphs are shown in Figures 1 and 2. They depict the continuous volume changes occurring in the infant's cry.

---

* Received in the Editorial Office on July 29, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study is supported by grant No. M-2875 of the National Institutes of Health, Division of Research Grants.

The *V-U* graph presents a clear picture of the onset of crying, the number of bursts, their relative intensity, and the pattern of ascent and descent. Since the graphic ammeter runs on a synchronous motor time is readily measured. Thus, given a stimulus to induce crying, the cry latency period (from stimulation to the onset of the cry) can be measured to a fraction of a second. The duration of each burst and the intervals of silence can be measured in the same way. The cry-sample shown in Figure 1 was obtained

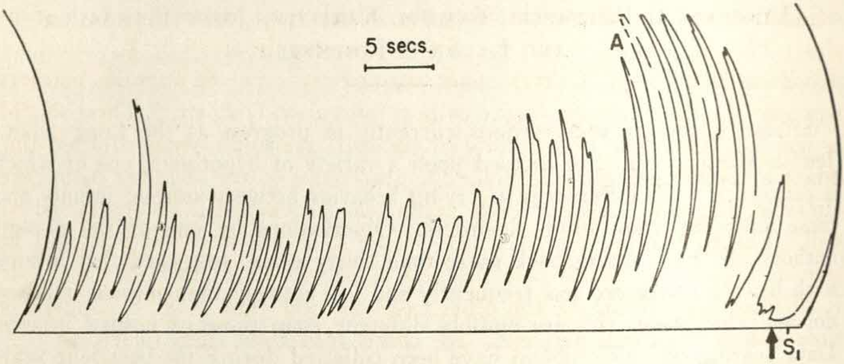


FIGURE 1

V–U graph of a cry-sample obtained from a normal girl, 3 days old. Graphs read from right to left.

S_1 = Point of stimulus.

Arrow = Onset of infant's vocal response.

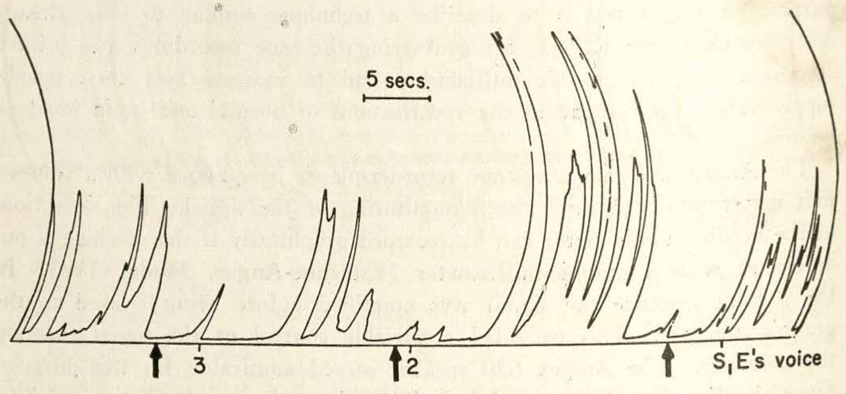


FIGURE 2

V–U graph of a cry-sample obtained from a mongoloid girl 2 days old.

S_1, 2, 3. = Points of stimulation.

Arrows = Onset of infant's vocal response.

E's = Experimenter's voice.

from a normal girl three days old. She had been stimulated at point S_1 (stimulation consisted in flicking the sole of the infant's foot once with the index finger) and the arrow indicates a vocal response 1.2 seconds later. Her cry is vigorous and sustained. No further stimulation was required. The sample shown in Figure 2 was taken from a mongoloid girl two days old. The vertical lines at the base of the chart show successive points of stimulation and the arrows indicate the initial vocal reaction to each stimulus. The cry latencies for the first three stimuli are 4.2, 1.2, and 3.4 seconds in that order. Curves appearing directly over the stimulus markers represent the experimenter's voice calling "stimulate 1, 2, etc." These should be ignored. Only the first minute of the full cry-sample obtained from this infant is shown in Figure 2. Since the gear ratios of the ammeter motor are alterable, the time scale for each graph may be expanded or contracted at will. The full cry-sample of which Figure 2 is only the first part will be seen in Figure 3. In this case the gear ratio was 45:45 and the time scale is 20 seconds per inch. In the other cases the gear ratio was 18:72 and the time scale is 5 seconds per inch. All charts read from right to left.

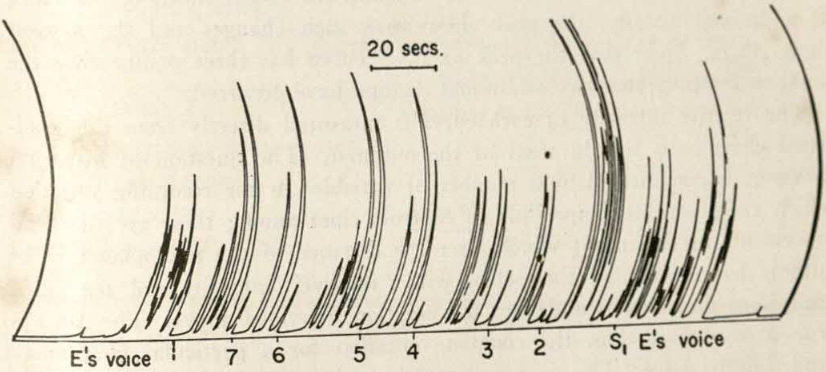


FIGURE 3

V–U graph of the complete cry-sample, the first part of which is shown in Figure 2.
S_1 and numbers 2 through 7 = points of separate stimuli.
E's = Experimenter's voice.

The graphs also reveal interesting patterns which reflect individual differences in crying behavior. The rapid and uniform bursts shown in Figure 1 are easily distinguished from the less active cry shown in Figure 2. The height of each curve is a measure of its intensity expressed in milliamperes. The ordinate scale is from zero to five milliamperes.

While a general inspection of the charts shows roughly the amount of crying activity, it also reveals that the individual curves differ in their form as well as their height: some of the curves are clean spikes, while others show "blips" of activity on the ascent, the peak, the descent, or combinations of the three. These differences combine with differences in duration and intensity to provide a unique quality for each curve. The clearly observable curve characteristics may be transformed numerically in the following manner: Each curve has an ascending period, a peak, and a descending period. In a clean spike the ascent is not interrupted by additional volumic activity, the peak is a clear point, and the descent shows no hesitation or upswing in volume. If the chronological order of the three periods is preserved; ascent, peak, and descent, the simple spike may be given the notation 0-0-0. No volume changes, in addition to the fundamental upswing and fall, occur during the ascent, at the peak or during the fall. When additional volumic activity does occur, the order of notation will indicate whether it occurred during ascent (first position), at peak (second position), or during descent (third position). Cardinal numbers may be used in these three positions to indicate the number of volume changes which appear. The notation 1-2-3 describes a curve in which the ascent shows one instance of additional activity, the peak shows two such changes and the descent shows three. Note that the peak of 1-2-3 curve has three points since the notation indicates that two additional changes have occurred.

The relative intensity of each curve is measured directly from the graduated chart as is the duration of the outburst. The question of intensity, however, is complicated by a number of variables in the recording situation which are practically impossible to control; chief among these are the head movements of the infant which alter the distance of the microphone. The authors have contented themselves with "relative" intensity and the value read from the chart is added to the notation described above. The same is true of duration: thus, the complete notation for a particular curve may read, 0-0-0-4.3-1.4: This is a spike of intensity 4.3 milliamps and lasting 1.4 seconds (see curve A in Figure 1.)

In a great many instances the curves run off the charts. In order to determine what the shape of the curve is, the tape is run again at a lower amplification. If the top of the curve is not revealed on the re-run, the tape is played at successively lower amplifications until the whole curve is seen and its notation recorded. Changes in amplification are never made during the run itself so that in a single record identical needle excursions represent the same input energy. In the work currently in progress at the Long Island Jewish Hospital, each tape is run through the ammeter at least twice.

Systematically, the amplifier is set at a reading of 5.0 and the tape is played. The amplifier is then set at 2.5 and the same tape is re-run. If lower or higher readings are required the tape is played again. All the while, audio-monitoring must go on simultaneously since the ammeter does not indicate what is cry and what is the experimenter's voice reading necessary information into the tape. Extraneous sounds must be crossed off the *V-U* chart as they are identified. Certain technical problems require further search, particularly the one concerning the relationship between frequency (cps) and "intensity" as measured in milliamps; but there is no doubt that the technique provides a permanent and easily measured description of the infant's cry. The pattern of the cry is plainly visible. The *V-U* charts have been especially useful in providing a means of measuring the temporal relationships within the cry. The results of these studies will be reported separately.

REFERENCES

1. KARELITZ, S., KARELITZ, RUTH AND ROSENFELD, LAURA S. Infants' vocalizations and their significance. *Mental Retardation: Proc. First International Med. Conf.*, 1960, (*ed.* by Bowman, P. W. and Mautner, H. V.) New York: Grune and Stratton, 439-446.
2. PETERSON, A. AND GROSS, E., JR. *Handbook of Noise Measurement* (4th ed.). New York: General Radio Co., 1960.
3. TIFFIN, J. Applications of pitch and intensity measurements of connected speech. *J. Acoust. Soc. Amer.*, 1934, 5, 225-234.

The Long Island Jewish Hospital
270-05 76th Avenue
New Hyde Park
Long Island, New York

Published as a separate and in *The Journal of Psychology*, 1961, 52, 429-444.

CONDITIONS FOR CREATIVE PROBLEM SOLVING*[1]

Department of Psychology, University of Michigan

L. RICHARD HOFFMAN

A. PURPOSE

Duncan (1959) concluded his review of recent research on problem solving by pointing to some major gaps in the area. One of the gaps is a lack of adequate theory due to "unanalyzed and non-dimensionalized variables" and to the "great variety of tasks to provide problems" for research. These are certainly justifiable complaints. Little progress has been made in problem solving theory since Dewey's description of the creative process: (1) the feeling of unease or dissatisfaction, (2) identification of the problem, (3) establishing hypotheses, (4) logical or empirical verification, (5) acceptance of the conclusion. This description offers a systematic approach to the creative process, but it lacks an explanation of when unease is felt, for example, or how hypotheses are generated; and thus tells us little about how creative problem solving might be encouraged.

The present paper is an attempt to fill part of the theoretical gap in problem solving with respect to a limited set of tasks. The framework to be used rests heavily on the creative thinking and research of Maier (1930, 1931, 1932), Duncker (1945), and Wertheimer (1945) in this area. Limiting the discussion to particular types of problems precludes any attempt to integrate all the various studies which have been called problem solving. Further, there are certainly conditions under which creative problem solving will occur other than those to be stated.

The purpose of the article is to present a framework capable of explaining the results of a number of studies, of suggesting next steps for research on the problem-solving process, and of pointing to the types of techniques which should be developed to enhance creative problem solving. The theory appears applicable equally to individual and group problem solving and even, possibly, to problem solving at the societal level.

---

* Received in the Editorial Office on July 31, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This paper was written as part of a project supported by USPHS research grant M-2704 from the National Institute of Mental Health, Public Health Service. The author expresses grateful appreciation to Norman R. F. Maier and Helen Peak for their constructive comments on the paper.

The paper will present the conditions for creative problem solving and illustrate their usefulness in explaining the results of some selected studies.

B. Creative Problem Solving: A Definition

Before stating the conditions for creative problem solving, a definition of the term is needed. Following Wertheimer (1945) and Maier (1933), creative problem solving is the process of forming a solution pattern which is a combination of previously dissociated experiences. These experiences may have been thought of previously as incompatible or irrelevant, or they may never have been considered together prior to the problem-solving situation. The new solution pattern is more than merely an abstraction from the elements of the "experiences," but is in fact a new organization of them, an organization which combines the elements of the experiences in a new and different way. In most cases the elements of the creative solution will bear different relations to each other than they did in their pre-solution organization. For example, Newton's gravitational principles put apples and cannonballs in a relation to each other they had never had before.

Despite the importance society attaches to creativity, the progress of research appears to require that the product of creative problem solving be evaluated presently only with reference to the person or group doing the problem solving. It is creative when it is new for him, regardless of how new it is for society. Anderson (1959, p. 261) has termed this type of definition "internal locus of evaluation" as distinct from "external evaluation." Bartlett's definition of thinking is also consistent with the present one: "the use of any contributory sources of evidence that are available to reach a terminal point which is treated as if it had not been achieved before . . ." (1958, p. 74).

Although these definitions of creative problem solving combine societal inventions with childhood discoveries, they seem fruitful in helping us learn about creativity in as many situations as possible, after which we might apply this knowledge to situations which are potentially societally creative. The definition also implies that everyone is creative or potentially so. And this is probably true, although we shall see from the conditions to be stated why most of us are not creative.

C. Conditions for Creative Problem Solving

The following conditions appear to be necessary for creative problem solving, as defined above.

Condition 1. *Differing, but comparable, cognitions must coexist.* The general term "cognition" is used here to indicate any type of ideation, such as the definition of a problem, the "direction" or approach to solving the problem, a solution to the problem, or even a fact or an interpretation of a fact. The cognitions should be comparable with respect to the aspect of the problem for which they are relevant, e.g., the definition of the problem or the interpretation of facts. Thus, while two different proposed solutions to a problem satisfy Condition 1, a solution to a problem and a definition of the problem do not. The former condition may result in a creative solution, but the latter will not unless an additional and different solution to, or definition of, the problem emerges. In this way the incongruity concept of Osgood, Suci and Tannenbaum (1957) is appropriate for defining "differing cognitions," as are similar concepts like Festinger's *dissonance* (1957), and Peak's *disparity* (1955, 1958), although at a somewhat more molecular level than is usually considered in problem-solving research. In terms of the above definition of problem solving, these differing cognitions would be the "previously dissociated experiences."

In a group setting, satisfaction of Condition 1 would entail at least two contributions concerning a problem, ordinarily from different group members. Such contributions commonly take the form of potential solutions. For example, a department head might be concerned that his department was not turning out production as fast as the demand for his product warranted. At a meeting of his staff, he might suggest the elimination of certain of the more expensive parts of the production operation in favor of a faster rate. A junior member of his staff might counter with the suggestion that quality should be maintained at all costs, and that, since quality takes time, the production rate will have to be sacrificed for quality considerations. Here are two differing cognitions, different solutions to a problem.

Condition 1 may be diagramed as in Figure 1 by two opposing arrows to indicate the differences between two ideas.

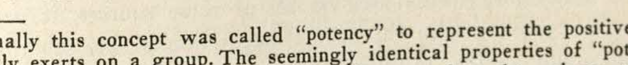


FIGURE 1

Valence. The representation of these two ideas by opposing lines foretells the concept of the valence of a cognition,[2] which must be introduced before going on to Condition 2. Valence is used here, analogously to Lewin's usage with respect to the attraction of goals (1935), to represent the degree of acceptibility a cognition may have for a problem-solving unit, i.e., the likelihood that it will be accepted by the unit, be it individual or group. Positive valence would indicate an acceptable cognition and negative valence an unacceptable cognition. A cognition may acquire positive valence from a variety of sources, and the effects of these sources may be presumed to accumulate to increase the valence of the cognition.

---

[2] Originally this concept was called "potency" to represent the positive force an idea usually exerts on a group. The seemingly identical properties of "potency" and Lewin's valence concept have caused the author to adopt the valence terminology in the interest of saving psychology one more idiosyncratic term.

A cognition may have positive valence for an individual as a function of the amount of his past learning about the subject, the frequency of his past experience with the cognition, the number of facts perceived to support the cognition, the degree to which the cognition fits in with other knowledge or attitudes he holds, the extent to which he "needs" to hold this cognition as a function of his personal motive structure, etc.

At the group level, the concept of valence may be translated to mean the degree to which a suggestion by a group member is acceptable to all members of the group. Although a solution will have different valence for each group member at any point in the discussion—even, presumably, at the time of final decision—since a group usually is asked for and produces a single solution, it is useful to think of the valence of any solution for the group as a whole. Group valence may turn out empirically to be merely some combination of individual valences, but this does not destroy the utility of considering it as a concept at the group level.

Substituting the word "group" for "individual" in the above listing of sources of positive valence, these sources also act at the group level. In addition, the following sources of positive valence in a group setting come to mind: the persuasiveness of the person presenting an idea, the total power of the person making the suggestion in the group, the enthusiasm with which the suggestion is made, the degree of support for the suggestion in the group, etc. For example, in the above illustration, the department head's suggestion has positive valence by virtue of his formal position in the organization. We may assume that his subordinate's suggestion would thereby have less positive valence. The department head's suggestion may acquire further positive valence if the rest of his subordinates are a group of "yes men," or if he seems to be enthusiastic about his suggestion, or if he was particularly persuasive in presenting his idea. If the valence of the department head's idea were based on all of these sources, it would then be so great as to overwhelm the counter-suggestion of his subordinate and the group would decide to eliminate the unnecessary operations. This situation is diagramed in Figure 2.

FIGURE 2

On the other hand, if the department head's suggestion were seen as a threat by the group, the resulting defensiveness might focus their support for the counter-suggestion, or the subordinate on his part might have been extremely persuasive, or facts might be marshalled to support the subordinate's proposal in opposition to the initial one. The force picture would thereby be changed (Figure 3) and the group would decide to maintain the present form of operation.

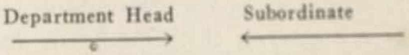


FIGURE 3

In other words, the valence of an idea acts as a force and, as in Lewin's model of conflict resolution (1935), the idea with the greatest positive valence is adopted. Of course, the valence of an idea is not a fixed value and may change during the problem-solving process, either increasing as new facts are offered to support the proposition or enthusiasm is generated for it, or decreasing as the support of the group moves to other alternatives.

If the valence of all alternatives remains small—nobody is particularly enthusiastic about any of the ideas—apathy may result and no decision may be reached. The existence of such conditions necessitates a minimum threshold concept, such that at least one alternative must acquire valence beyond this minimum for a group to arrive at any decision. When all alternatives have low valence, the mere suggestion, "Let's do 'A' so we take some kind of action," will give "A" enough positive valence for the group to agree to take that action. It is probable, however, that the type of acceptance, i.e., willingness and enthusiasm to carry out the decision, usually posited to accompany the process of group decision (e.g., Maier, 1952; Coch and French, 1948), can be achieved only when the valence of the alternative finally agreed upon is considerably greater than the minimum threshold level, and/or the solution has substantially higher than minimum valence for all the group members. Such a high value would reflect the group's identification with and enthusiasm and support for the agreed-upon decision.

The elaboration of the valence concept and the description of its mode of operation is necessary as a prelude to the second condition necessary for creative problem solving.

Condition 2. *At least two differing cognitions must acquire greater positive valence than the minimum threshold value* (but less than some maximum value which will be discussed in Condition 3) *and must be approximately equal in value, so that none of the alternatives can be accepted and an impasse is reached.* The rationale for the desirability of an impasse between cognitions rests on the assumption that if two cognitions are equally desirable, yet neither is acceptable, there are advantages and disadvantages regarding, or facts supporting, each. The creative solution to such a problem would exploit the advantages of both alternatives and minimize the disadvantages, or would incorporate the seemingly incompatible facts supporting the several alternatives. To achieve such a solution would require the type of restructuring that Wertheimer has discussed (1945).

Simpson (1960) has provided us with an example of such a creative solution in the realm of evolutionary theory. He points to Neo-Darwinism, the natural selection process, and Mutationism as the current fact-supported theories of evolution. The two theories, however, have been offered as contradictory explanations of the process of evolution. Neo-Darwinism proposes that evolution is an adaptive selective process, while Mutationism suggests that adaptive changes occur randomly with the non-adaptive mutants dying off. Simpson concludes that "Mutationism is not an alternative to Neo-Darwinism but a supplement of it . . . if the actual course of evolution is to a large extent adaptive, then some additional factor or process must frequently intervene between the occurrence of mutations and the incorporation of some of them into evolving populations . . . . Natural selection is just a process, and . . . is in large part a synthesis of selection theory and mutation theory." Thus, opposing fact-supported theories are reconciled into a single more eleborate theory which encompasses both sets of facts.

Impasses, however, have been known to cause breakdowns in problem-solving conferences. Pre-contract negotiation conferences between labor and management frequently result in impasses, followed by collapse of negotiations and strikes.

The three problem-solving situations described above—the single prepotent cognition, the apathetic, and the impasse—are diagramed in Figure 4 with their probable outcomes. Two possible outcomes are indicated for the impasse condition, no solution or a creative solution. A third possibility is for the situation to regress to the single prepotent cognition state. To account for the emergence of a creative solution rather than a breakdown of problem solving in the impasse situation a third condition is necessary.

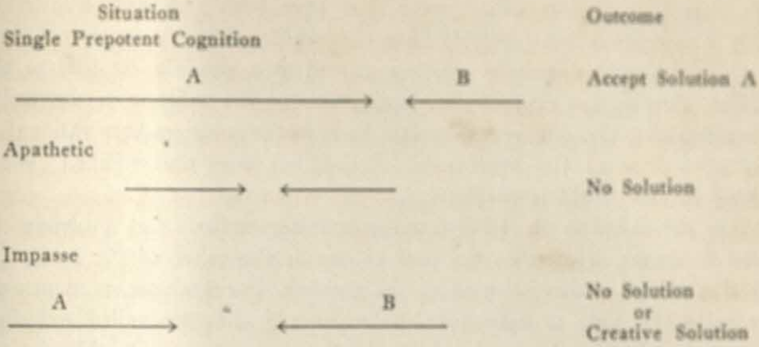


FIGURE 4
PROBLEM-SOLVING SITUATIONS AND THEIR OUTCOMES

Condition 3. *Problem solving must occur in a situation in which the problem-solving unit is required to arrive at the best possible decision.* This requirement may be achieved either through externally imposed requirements, e.g., by a company president's demand for the best solution from his subordinates, or from the needs of the group members themselves, e.g., to show everyone how bright they are. The requirement to arrive at the best possible decision may be considered as a boundary condition for the problem-solving unit and must therefore be more powerful than the valence of any of the ideas offered. Diagrammatically, Lewin's life space model may be borrowed again to represent this boundary around the problem-solving situation (see Figure 5). This boundary must be of sufficient strength to prevent the problem-solving unit from "leaving the field." Condition 3 does not permit the unit to "agree to disagree."

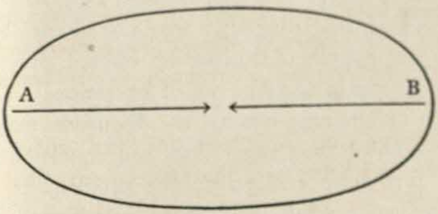


FIGURE 5

The presence of a requirement to make the best decision may also help to satisfy Conditions 1 and 2. Motivation for the best solution should encourage the production of ideas and prevent the premature acceptance of any early suggestions before other possibilities have been explored. Failure to satisfy Condition 3 has probably been the source of many people's failure to solve problems adequately. Having arrived at a possible solution to the problem, they do not explore the ramifications of the solution to determine its inadequacies. Or, discovering certain flaws in the solution, they rationalize their acceptance of the inadequate solution by some phrase like, "Well, nothing in this world is perfect."

Many breakdowns in labor-management negotiation can probably be traced to a lack of need on the part of one or the other of the parties to arrive at a high quality solution to the problem. For example, management may feel it is able to tolerate a strike since it has stockpiled sufficient inventory to supply the needs of its customers for a considerable period. Or labor may recognize that seasonal layoffs are about to occur anyway and they have nothing to lose by striking.

The postulation in Condition 3 of the existence of a force field toward arriving at a best solution places an upper limit on the valence of any of the alternatives being considered. Where the valence of any of the alternatives exceeds this upper bound, it must be reduced if a creative solution is to emerge. Otherwise this highly valent solution will be accepted before the creative solution is discovered. Often, in practical group situations, such high valence reflects extreme emotional identification with the idea on the part of one of the parties, for example, defensive behavior. If the emotional support for such a proposal is reduced, for example by the Two Column Method (Maier, Solem, and Maier, 1957), a balance of valences frequently can be achieved at a manageable level—i.e., in the range between the minimum threshold level and the level of the force towards solution. The problem can then be restructured and a creative solution developed. The need to arrive at a best decision may help a group member to recognize that his own first suggestion may not always be the best one.

For example, in the hypothetical situation offered previously, where the department head suggested a speeding up of operations and one of the subordinates insisted on maintaining quality, the discussion may have progressed to the point where each of these suggestions had equal valence. The department head should then have reidentified the problem as "how to achieve a higher rate of production while maintaining an adequate level of quality." The solution to such a problem would probably result in the elimination of certain parts of the present production process as well as suggestions for its reorganization. It might be possible in this way to achieve the desired level of production while maintaining the same standards of quality.

The ability of the department head to so rephrase the conflicting alternatives illustrates a fourth condition desirable for creative problem solving.

Condition 4. *The cognitive components of the valence of each of the conflicting alternatives should be abstracted and their points of conflict recognized.* Although Maier (1931) has shown that a subject need not be able to report the basis for his solution in order to achieve it, awareness of the conflict to be resolved is probably helpful in arriving at a solution. By becoming aware of the conflicting elements of the problem, the subject can adopt a fruitful direction for solving the problem and he is in a position to search for variations in these elements which are appropriate for the solution (cf., Wertheimer's "recentering," 1945).

The inability to define the cognitive basis of an impasse situation often defeats the search for a creative solution. In groups, opposing solutions may become inseparable from personal dislike between the people suggesting the solutions, so that a real cognitive conflict is lost by its being seen only as an emotional conflict. Such a situation may deteriorate to the point where the motivation to arrive at a best solution is superseded by motivation to resolve the conflict in the easiest way and lead to the adoption of an inferior solution. Individuals, too, may become discouraged by their inability to determine just what relationship two ideas have to each other and, therefore, accept one alternative in order to escape an uncomfortable situation.

In summary, if the first three conditions of this paper are satisfied, the stage appears to be set for creative problem solving to occur. The problem-solving situation at that point has the following characteristics.

At least two alternative cognitions (e.g. solutions) have approximately equal positive valence which is sufficient to make either one acceptable. Because their valences are equal, however, neither can be accepted in preference to the other. The problem-solving unit (individual or group) has sufficient need to produce the best possible solution so that he will neither withdraw from the problem nor accept one of the alternatives for the sake of expediency (or for some other reason). Beyond the satisfaction of Conditions 1 through 3, a fourth condition has been postulated as desirable, although not necessary for achieving a creative solution. Condition 4 states that when the problem solver is in the situation just described—where Conditions 1, 2, and 3 have been satisfied—he will be more likely to achieve a creative solution if he is able to identify the conflicting cognitions present.

D. THE RESEARCH EVIDENCE

Some examples from the research literature will illustrate the usefulness of viewing problem solving in terms of these conditions. Unfortunately for the purpose of the review, problem-solving research has been concerned mostly with conditions which prevent successful problem solving, and little with conditions which increase problem-solving effectiveness. Thus, much evidence could be marshalled to show that failure to satisfy the postulated conditions will interfere with creative problem solving, but little evidence is available to substantiate the productive effects of satisfying the conditions. Also, most of the tasks used in these studies do not completely satisfy the definition of creative problem solving being used here, since satisfaction of Conditions 1 and 2 is often sufficient for the subjects to attain the solution. Even in these tasks, however, failure to solve the problem or to arrive at the best solution may be due to poor motivation of the subjects in the problem situation, thus not satisfying Condition 3.

Luchins' series of "Einstellung" experiments (1942) using water-bottle problems are interesting in light of the stated conditions. He showed that past experience with a particular method for solving problems caused subjects to use that method in solving problems which could have been solved using a simpler method, and reduced their ability to solve similar problems for which the method was inappropriate. The past experience seemingly blinded subjects to the existence of an alternative method, thus failing to satisfy Condition 1 (differing cognitions), and the "Einstellung" method has sufficient positive valence to be accepted, thus failing to satisfy Condition 2 (two equally valent alternatives). Furthermore, most subjects interpreted the experimenter's instructions to find the "best solution" to mean an accurate answer, not necessarily the simplest or most direct answer. Thus Condition 3 was not satisfied in most of the experiments. When subjects were given the problems untimed, instructed to look over their work, told to find "a more sensible and *more direct* solution to the problem," and given experience with a problem having another method for solution, thus presumably satisfying the first three conditions of this paper, recovery from the Einstellung effect was "very good" (Luchins, 1942, p. 59, Experiment 5).

Duncker's (1945) classical studies of "functional fixedness" illustrate a more specific Einstellung situation which fails to satisfy the conditions. The valence of the object's most common function, especially in the "pre-utilization conditions," is so great that it surpasses the lower threshold value and is accepted as the sole use of the object. The valence of other functions of the object is presumably below threshold, so the problem cannot be solved.

Flavell, Cooper and Loiselle's report (1958) shows that the valence of functions other than the "fixed" one can be increased by giving subjects a variety of "non-specialized" pre-utilization experiences with the object. Their findings suggest that Condition 1, at least, was satisfied and resulted in a higher than ordinary proportion of successful problem solutions.

Studies by Rokeach (1960) and Milton (1958) showing that subjects' attitudes may interfere with effective problem solving, may be interpreted as evidence that certain attitudes interfere with the emergence of differing perceptions or orientations to problems, thus failing to satisfy Condition 1. Carey's (1958) ability to improve women's problem-solving success merely by changing their attitudes toward problem solving, tends to support this interpretation.

Hoffman's (1959) and Hoffman and Maier's (in press) studies of the effects of personality homogeneity and heterogeneity on group problem solving explicitly attempted to satisfy Condition 1 in groups. They assumed that people with dissimilar personalities would have a greater variety of perspectives on problems than would people with similar personalities and, therefore, have greater problem-solving success. The consistent problem-solving superiority of the heterogeneous groups confirmed this hypothesis. The fact that none of the other conditions of this paper was controlled in these studies, however, probably accounts for the small differences obtained.

Studies by Torrance (1955), Riecken (1958) and Maier and Solem (1952) show how formal status, "talkativeness," and majority opinion, respectively, may contribute to the valence of incorrect solutions and prevent the acceptance of correct answers—failure to satisfy Condition 2.

Torrance showed that in Air Force crews solving Maier's Horse Trading Problem, the status hierarchy was related directly both to the degree to which groups accepted the correct solution held by one of its members, and to the amount of talking by each of the group members. Solutions suggested by low ranking members rarely acquired enough valence to be accepted by the group.

Riecken's (1958) study of "talkativeness" parallels Torrance's findings in showing that the most talkative group member, regardless of the adequacy of his information, exerts considerable influence over his group's decision. Riecken compared the effect of giving a hint concerning the "elegant" (creative) solution to the most talkative versus the least talkative members of groups. The hint was much more often accepted from the most talkative member. Furthermore, Riecken's data suggest that "success in gaining acceptance for the elegant solution (for both high and low talkers) seems to depend, not on amount of opposition aroused, but on amount of support the hint-holder musters," thus offering direct support for the present valence analysis.

Maier and Solem (1952) have shown that a discussion leader may aid in overcoming majority influence. In most of the unled groups, the initial solution held by the majority was retained after group discussion and, where the majority was initially incorrect, few members changed to the correct answer as a result of the discussion. The effect of the discussion leader in helping the group arrive at the correct answer was most marked in those groups where the majority initially had the wrong answer. Maier and Solem interpreted their findings as revealing the discussion leader's effective use of minority opinions (raising their valence) to overcome the influence of the majority.

Another series of studies may be interpreted to show that satisfaction of the first three conditions yields creative problem solutions. These studies used a role-playing problem in which two alternative solutions are favored by people differing in status position and authority. Maier and Hoffman (1960) showed that groups that solved the problem twice produced a higher proportion of integrative (creative) solutions the second time than the first. Hoffman, Price and Maier (unpublished data) later showed that this effect was produced in part by an increase in the influence of the low status group members over the second solution. Presumably the valence of the alternative favored by the low status members was increased the second time, more nearly equalled the valence of the other alternative, and caused the group to develop a creative solution. Even where the valence of each alternative was high originally, the second problem-solving session permitted other solutions to acquire greater valence as the group was forced to reexamine the problem situation. In a later study using the same problem, Hoffman, Harburg and Maier (in press) experimentally increased the valence of the low status members' favored alternative and produced a high proportion of integrative solutions. When conflicting solutions which had approximately equal valence were put into opposition, groups searched for new solutions which incorporated the advantages of each of the original solutions.

Finally, two studies may be cited to show very different reasons why

Condition 3 (need to arrive at a best decision) may not be satisfied. Typical of the experimental—and probably practical—effects of time pressure is to make "the best decision" a less desirable goal and to allow the subject to be satisfied with a less adequate, sometimes inadequate, solution. In one variation of his Einstellung experiments, Luchins (1942) showed that, when he made the problems a "speed test," the Einstellung effect was greater than in the general study; many subjects submitted the Einstellung solution to the "critical" problem (for which it was inappropriate), and a number of subjects gave up.

However, the need postulated in Condition 3 may be unsatisfied, not only because of stronger conflicting needs, but also because the situation is not conducive to such motivation. The studies of business decision making by Cyert, Dill and March (1958) show how the presence of "organizational slack"—a surplus of resources—allowed certain decisions to be accepted without either adequate criticism, consideration of other alternatives, or even a definition of the problem to be solved. Motivation to produce a "best decision" (Condition 3) was low due to the feeling that the costs of a poor decision could be absorbed by the surplus resources available. Thus failures to satisfy Condition 3 caused failures of Conditions 1 and 2 and resulted in poor decisions.

E. IMPLICATIONS AND CONCLUSIONS

Although the evidence presented is certainly not conclusive, these formulations of the necessary conditions seem useful for understanding the results of a variety of studies of problem solving, both by individuals and groups. If satisfaction of these conditions is conducive to creative problem solving, we can now look for behaviors and attitudes which will tend to satisfy the conditions.

Techniques need to be developed to promote these conditions for both individuals and groups. A few successful attempts have been made in this direction. Maier (1933) increased the proportion of successful problem solutions by lecturing subjects to seek for various problem difficulties and to "keep your mind open for new meanings (and) . . . for new combinations." Maltzman and his associates (Maltzman, Brooks, Bogartz and Summers, 1958; Maltzman, Bogartz and Breger, 1958; Maltzman, Simon, Rasken and Licht, 1960) have been orienting subjects toward the same objectives by forcing them to generate unusual (low valence) associations to common stimuli. These methods appear conducive to establishing Condition 1.

Shaw's (1960) results suggest that subjects' motivation to achieve a high quality product (Condition 3) may be increased by increasing their share of responsibility for the group output.

Maier's human relations training course includes various functions a group discussion leader may perform to help assure the satisfaction of these conditions (Maier, 1952, 1958). Among these functions are: (*a*) "respect minority opinions"—promotes Condition 1 by revealing differing viewpoints; (*b*) separate the idea *getting* process from the idea *evaluating* process—promotes Condition 2 by allowing all ideas initially to achieve more than minimum positive valence without revealing their possible negative valence; (*c*) "state the problem in such a manner that there is mutual interest"—promotes Condition 3 by assuring that a good solution to the problem is desired by all members of the group.

Research is now under way to operationalize the concepts defining the conditions set forth in this paper and to test the adequacy of the formulations of the conditions. The evidence cited suggests that such effort may be rewarded in giving psychologists greater understanding of the problem-solving process and greater success in producing more creative solutions to problems.

REFERENCES

1. ANDERSON, H. H. Creativity in perspective. *In* H. H. Anderson (*Ed.*), *Creativity and its Cultivation.* New York: Harpers, 1959.
2. BARTLETT, F. Thinking. New York: Basic Books, 1958.
3. CAREY, GLORIA L. Sex differences in problem-solving performance as a function of attitude differences. *J. Abnorm. & Soc. Psychol.*, 1958, **56**, 256-260.
4. COCH, L., & FRENCH, J. R. P. Overcoming resistance to change. *Hum. Relat.*, 1948, **1**, 512-532.
5. CYERT, R. M., DILL, W. R., & MARCH, J. G. The role of expectations in business decision making. *Admin. Sci. Quart.*, 1958, **3**, 307-340.
6. DUNCAN, C. P. Recent research on human problem solving. *Psychol. Bull.*, 1959, **56**, 397-429.
7. DUNCKER, K. On problem solving. *Psychol. Monogr.*, 1945, **58**, No. 5 (Whole No. 270).
8. FESTINGER, L. A Theory of Cognitive Dissonance. Evanston, Ill.: Row, Peterson, 1957.
9. FLAVELL, J. H., COOPER, A., & LOISELLE, R. H. Effect of the number of pre-utilization functions on functional fixedness in problem solving. *Psychol. Rep.*, 1958, **4**, 343-350.
10. HOFFMAN, L. R. Homogeneity of member personality and its effect on group problem solving. *J. Abnorm. & Soc. Psychol.*, 1959, **58**, 27-32.
11. HOFFMAN, L. R., HARBURG, E., & MAIER, N. R. F. Differences and disagreement as factors in creative group problem solving. *J. Abnorm. & Soc. Psychol.*, in press.

12. Hoffman, L. R., & Maier, N. R. F. Quality and acceptance of problem solutions by members of homogeneous and heterogeneous groups. *J. Abnorm. & Soc. Psychol.*, (in press).

13. Lewin, K. A Dynamic Theory of Personality. New York: McGraw-Hill, 1935.

14. Luchins, A. S. Mechanization in problem solving—the effect of *einstellung*. *Psychol. Monogr.*, 1942, **54**, No. 6 (Whole No. 248).

15. Maier, N. R. F. Reasoning in humans. I. On direction. *J. Comp. Psychol.*, 1930, **10**, 115-143.

16. ———. Reasoning in humans. II. The solution of a problem and its appearance in consciousness. *J. Comp. Psychol.*, 1931, **12**, 181-194.

17. ———. An aspect of human reasoning. *Br. J. Psychol.* (Gen. Sect.), 1933, **24**, 144-155.

18. ———. Principles of Human Relations. New York: Wiley, 1952.

19. ———. The Appraisal Interview: Objectives, Methods and Skills. New York: Wiley, 1958.

20. ———. Maier's Law. *Amer. Psychologist*, 1960, **15**, 208-212.

21. Maier, N. R. F., & Hoffman, L. R. Quality of first and second solutions in group problem solving. *J. Appl. Psychol.*, 1960, **44**, 278-283.

22. Maier, N. R. F., & Solem, A. R. The contribution of a discussion leader to the quality of group thinking. *Hum. Relat.*, 1952, **5**, 277-288.

23. Maier, N. R. F., Solem, A. R., & Maier, A. Supervisory and Executive Development. New York: Wiley, 1957.

24. Maltzman, I., Bogartz, W., & Breger, L. A procedure for increasing word association originality and its transfer effects. *J. Exp. Psychol.*, 1958, **56**, 392-398.

25. Maltzman, I., Brooks, L. O., Bogartz, W., & Summers, S. S. The facilitation of problem solving by prior exposure to uncommon responses. *J. Exp. Psychol.*, 1958, **56**, 399-406.

26. Maltzman, I., Simon, S., Raskin, D., & Licht, R. Experimental studies in the training of originality. *Psychol. Monogr.*, 1960, **74**, No. 6 (Whole No. 493).

27. Milton, G. A. Five studies of the relation between sex-role identification and achievement in problem solving. New Haven: Dept. Industr. Admin. & Dept. Psychol., Yale Univer., ONR Tech. Rep. No. 3 (Contract Nonr 609-20), 1958.

28. Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. The Measurement of Meaning. Urbana, Ill.: University of Illinois Press, 1957.

29. Peak, Helen. Attitude and motivation. *In* M. R. Jones (*Ed.*), *Nebraska Symposium on Motivation*, 1955. Lincoln: University Nebraska Press, 1955.

30. ———. Psychological structure and psychological activity. *Psychol. Rev.*, 1958, **65**, 325-347.

31. Riecken, H. W. The effect of talkativeness on ability to influence group solutions of problems. *Sociometry*, 1958, **21**, 309-321.

32. Rokeach, M. The Open and Closed Mind. New York: Basic Books, 1960.

33. Shaw, D. M. Size of share in task and motivation in work groups. *Sociometry*, 1960, **23**, 203-208.

34. Simpson, G. G. The world into which Darwin led us. *Science*, 1960, **131**, 966-974.

35. Torrance, E. P. Some consequences of power differences on decision making in permanent and temporary three-man groups. *In* Hare, A. P., Borgatta, E. F., & Bales, R. F. (*Eds.*), *Small groups.* New York: Knopf, 1955.
36. Wertheimer, M. Productive Thinking. New York: Harper, 1945.

Department of Psychology
University of Michigan
Ann Arbor, Michigan

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 445-455.

LYSERGIC ACID DIETHYLAMIDE (LSD-25) XXXIII. EFFECT ON SURFACING BEHAVIOR OF THE CARP AND POSSIBLE APPLICATION TO GAME CONSERVATION AND SURVIVAL KITS*[1]

Long Island Biological Association, Cold Spring Harbor, N.Y., and State Hospital, Central Islip, N.Y.

H. A. ABRAMSON, H. H. GETTNER, M. P. HEWITT, AND G. DEAN

Howard A. Loeb (8, 9) has stated that, "None of 1,800 assorted compounds force-fed to carp has produced more than random surfacing or any directed movements of the fish." Any class of compounds which produces surfacing and directed movements of carp, therefore, may be of importance in the control of fish populations, as well as in the study of the physiology of the fish themselves. It is the purpose of this communication to show that a group of compounds related to lysergic acid diethylamide (LSD-25) produces surfacing behavior of carp with the movement, therefore, directed toward the surface, providing the pharmacologic effect that has been sought through forced feeding. Our experiments have not as yet included feeding these drugs to fish, although extensive work on man (2) would indicate that the material is absorbed from the intestinal tract. The drugs are active when given intraperitoneally to the Siamese fighting fish (3).

It has been previously shown in work from this Laboratory that very small quantities of derivatives of d-lysergic acid, like lysergic acid diethylamide (LSD-25) and lysergic acid ethylamide (LAE-32) have an important effect on the Siamese fighting fish (1, 5). Less than 0.5 mcg per ml of LSD-25 in the outside liquid causes most of the fish to go to the surface of the liquid and remain relatively immobile. Figure 1 illustrates, in somewhat less than actual size, three juvenile forms of the Siamese fighting fish at the surface of a small volume of liquid containing 5 mcg of LSD-25 per ml. The nose up-tail down position of the fish is shown. At this concentration the fish may remain stuporous and at the surface for

---

* Received in the Editorial Office on July 31, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was supported in part by a grant from Sandoz Pharmaceuticals, Hanover, New Jersey.

several days. All of the fish usually recover from treatment with this type of lysergic acid derivative.

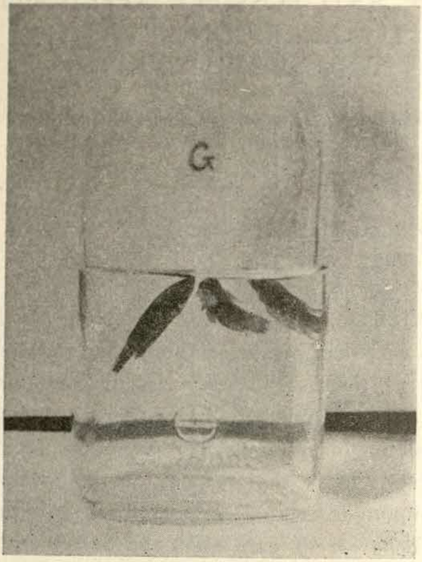


FIGURE 1
Three juvenile forms of the Siamese fighting fish at the surface of a small volume of liquid containing 5 mcg of LSD-25 per ml. The nose up-tail down position of the fish is shown. The fish are less than 2 inches long.

The need for improved sampling methods of fish collection in lakes is well brought out by the recent papers of Loeb (6, 7, 8), who describes both electrical and rotenone methods for capture and recapture of fish in small lakes from 30 to 300 acres in size. It is of importance to investigate the action of compounds like lysergic acid diethylamide on fresh-water fish in larger bodies of water, in connection with game conservation programs and the possibility of employing drugs of this type in survival kits. This paper is a preliminary report on laboratory trials which extend our earlier experiments to the carp. Larger volumes of liquid were employed, and much larger fish than hitherto were investigated.

A. Experimental

1. *Effect of LSD-25 on Goldfish*

The first experiments were performed on small goldfish commonly available in pet shops. In our earlier experiments it was found that goldfish respond to LSD-25 in somewhat the same way that Siamese fighting fish do. The reaction of goldfish was not so readily observed in small volumes of liquid (100 ml for 10 fish). For this reason a small aquarium was used to study goldfish. A 6-liter tank was found suitable for studies of this type. Note in Figure 2 the effect of 2 mcg per ml of LSD-25 on goldfish observed in two 6-liter tanks. Tank 1 is the control tank to which no LSD-25 or other chemical was added. The darkish forms at the bottom of Tank 1 are clusters of 10 goldfish. In the absence of disturbing influences the fish rarely rise to the surface during the course of the experiment. Tank 2, also containing 10 goldfish of the same size, strikingly illustrates the effect of 2 mcg per ml of LSD-25 on their surfacing behavior. After 5 minutes movement of the fish is readily observed with surfacing behavior almost complete after 10 minutes. This effect is very similar to the phenomena observed on the Siamese fighting fish. The surfacing behavior also occurs when a slow stream of air is bubbled through the tank. It should be noted that streaming air bubbles stimulate the fish. That is, the bubbles of air do not destroy the compound very rapidly. For this reason 2 minutes should elapse after turning off the air stream before a count is made of the fish at the surface. The size of the small goldfish may be readily estimated from the tank height which is 8 inches.

2. *Effect of LSD-25 on Large Carp*

Through the kindness of Dr. Howard A. Loeb of the New York State Conservation Department, carp at least 2 pounds in weight were sent to this Laboratory for study. The fish were kept in a running spring close to the Laboratory. Compare Figure 3, an experiment on these large carp in cold winter tap water, with an experiment on the much smaller species of carp, the goldfish (Figure 2). Each cylindrical glass tank in Figure 3 contains 44 liters of solution. The size of the fish may be estimated from the tank diameter which is 15.5 inches. Note in Figure 3 that Tank 2 rather than Tank 1 is the control tank. During the course of an hour, throughout which air was bubbled through both tanks, the control fish in Tank 2 remained at the bottom of the tank. However, the fish in Tank 1, to which 2 mcg per ml of LSD-25 were added, had reached the surface of the liquid in the tank within 45 minutes. They remained there in very much the same way as the Siamese fighting fish and goldfish (see Figures 1 and 2).

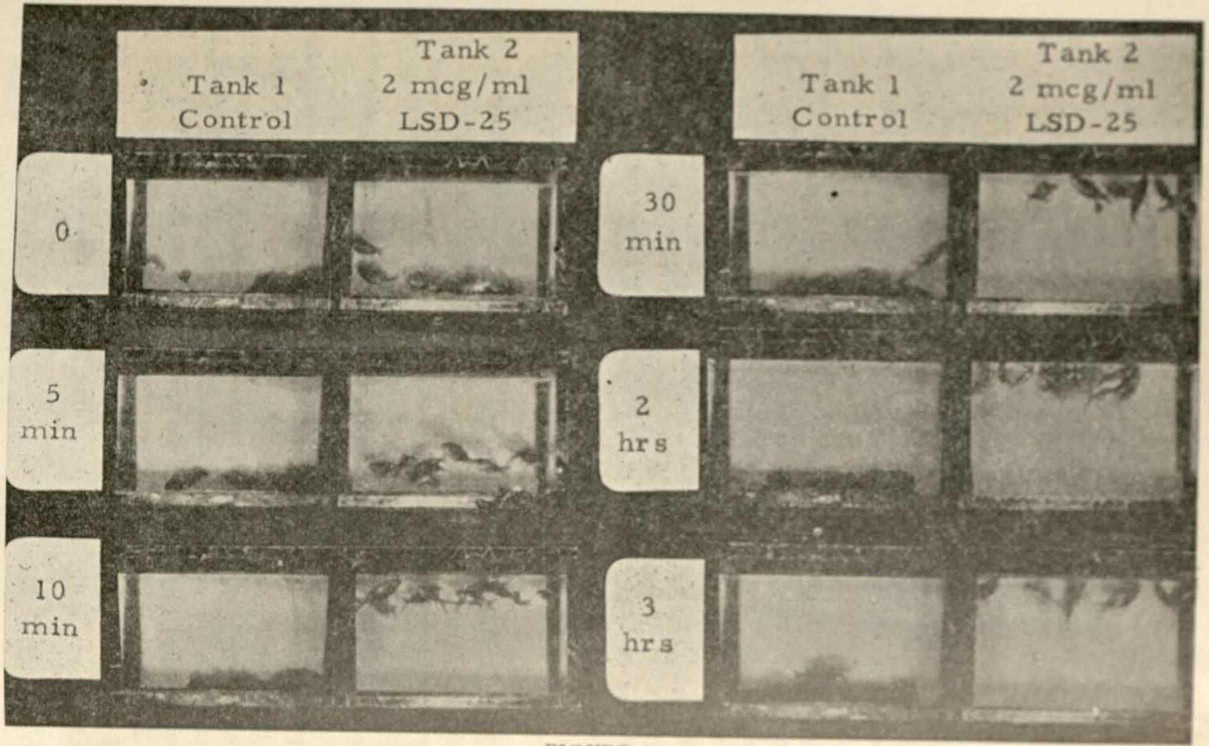


FIGURE 2

The effect of 2 mcg per ml of LSD-25 on goldfish observed in two 6-liter tanks. Tank 1 is the control tank to which no LSD-25 or other chemical was added. The darkish forms at the bottom of Tank 1 are clusters of 10 goldfish. Tank 2 also containing 10 goldfish of the same size, strikingly illustrates the effect of 2 mcg per ml of LSD-25 on their surfacing behavior. Note that the effect of LSD-25 is visible within 5 minutes and complete within 10 minutes under experimental conditions.

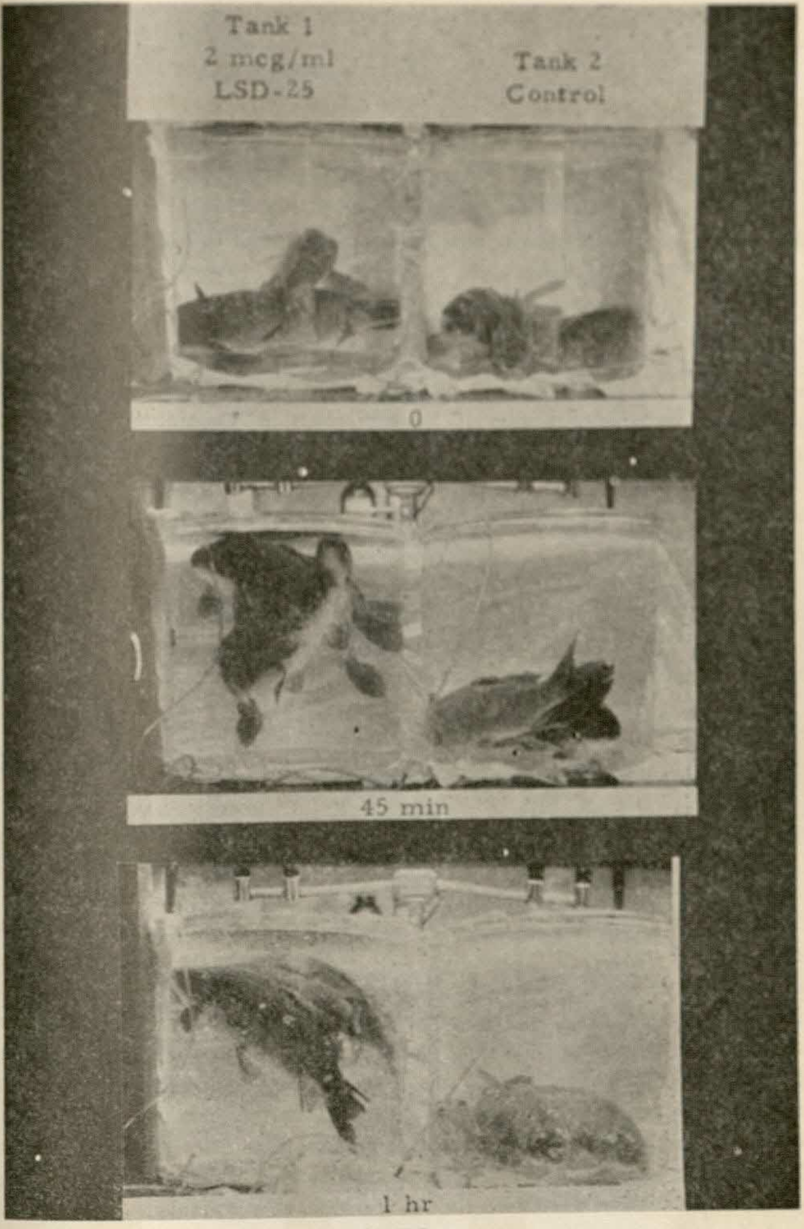


FIGURE 3

The effect of LSD-25 on large carp. Tank 2 is the control tank. Typical surfacing behavior due to LSD-25 is shown by the fish in Tank 1.

B. PROTOCOL OF TYPICAL EXPERIMENTS

Date: February 23, 1961

Control Tank: (Capacity 20 liters) Contained 3 carp approximately 2 pounds each.

Experimental Tank: (Capacity 20 liters) Contained 4 carp approximately 2 pounds each.

Air was bubbled through each tank. In the control tank, observed for 1 hour and 40 minutes, the fish stayed very quietly at the bottom. In the experimental tank, to which 40 mg of LSD-25 were added, the fish showed some signs of reacting within 10 minutes. Within 30 minutes they started to show typical surfacing behavior. At 30 minutes an additional 20 mg of LSD-25 were added. At 50 minutes all the carp in the experimental tank were at the surface and stayed there until removed from the tank at 3½ hours.

In order to test the reactivity of the control fish 40 mg of LSD-25 were added to the control tank at 1 hour and 40 minutes. Ten minutes later they began to show a typical LSD-25 reaction which was completed within 20 minutes probably because of the higher temperature of the water.

Date: March 2, 1961

Control Tank: (Capacity 44 liters) Contained 3 carp approximately 2 pounds each.

Experimental Tank: (Capacity 44 liters) Contained 3 carp approximately 2 pounds each.

At 12:30 p.m. 88 mg of LSD-25 were dissolved in the water of the experimental tank. After the fish had been exposed to LSD-25 for 10 minutes they showed signs of LSD-25 activity. Respiration was apparently increased and the mouth movements of the drugged fish were different from those of the control fish. At 30 minutes all 3 fish in the tank containing LSD-25 were at the surface of the liquid in a nose up-tail down position, with the control fish showing some restlessness but remaining at the bottom of their tank. For the subsequent hour the fish in the tank containing LSD-25 remained at the surface, from time to time moving and even swimming backwards. Of course, they were somewhat cramped in the round tanks and their movements were governed, perhaps in part, by their difficulty in maneuvering for space. It was extremely difficult for all three fish to be completely at the surface at one time.

The fish were returned to our outside running water pool after 1½ hours in the tank containing LSD-25. In the pool they continued to stay at the surface, moving around but not going to the bottom at all. Two hours after these 3 fish were placed in the pool they were still at the surface. Normal carp cluster at the bottom even when food is introduced. The fish under LSD-25 ignored food scattered in the pool.

Through the kindness of Mr. Edward White of the New York State Fish Hatchery at Cold Spring Harbor an experiment was done to estimate if eastern brook trout reacted similarly as carp to LSD-25. This single experiment is reported because of the possible importance of LSD-25 on brook trout in conservation programs.

1. *Effect of LSD-25 on Brook Trout*

Eight Salvelinus fontinalis (eastern brook trout) were placed in a large vessel (30 liters) of artesian well water at 11:55 a.m. Shortly before this the fish had been fed. At 12 noon, 40 mg of LSD-25 were added. The fish varied in size from 5 to 10 inches and were approximately 1 year old. At 12:15 p.m. the fish vomited. Most of the water was drawn off. New water was added and the 40 mg of LSD-25 were replaced. The fish vomited again at 12:20 p.m. At 12:25 an additional 20 mg of LSD-25 were added. The swimming function seemed to be impaired in some of the fish, and they remained on the bottom of the vessel until the effect of the LSD-25 apparently decreased. A few fish went to the surface of the water but did not remain there. Their reaction was not similar to the carp.

At 1:10 p.m. when the experiment was discontinued all fish were swimming slowly around in the vessel. They appeared to be somewhat drugged, but not enough to interfere with essentially normal behavior. Mr. White thought their behavior might represent oxygen deprivation.

All of the trout survived treatment with this dosage of LSD-25.

2. *Effect of LAE-32 on the Carp*

Figure 4 illustrates the effect of LAE-32 on small goldfish ½ hour after 2 mcg per ml of LAE-32 were added to their 6-liter tank. Air was bubbled through both tanks. The goldfish, stimulated by the air bubbling, show more activity in the control tank on the left than in the absence of bubbling.

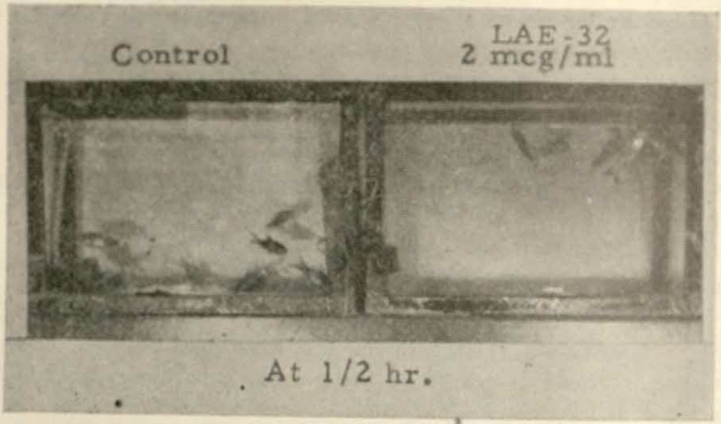


FIGURE 4
The effect of LAE-32 on small goldfish ½ hour after 2 mcg per ml of LAE-32 were added to their 6-liter tank. Note the surfacing behavior.

However, the fish in the right tank treated with LAE-32 show an effect identical with LSD-25. The similarity of the effect of LAE to LSD, however, was pointed out in one of our earlier papers on the Siamese fighting fish (5). The larger carp were next tested with LAE-32. Figure 5 shows that 2 mcg per ml of LAE-32 also affect the larger fish, qualitatively speaking, in the same way as LSD-25. There is a difference, however, in the rate at which LAE-32 reacts. LAE-32 takes longer to act and the larger carp do not remain as consistently at the surface as they do with a similar concentration of LSD-25.

Experiments with LAE-32 are significant for the following reasons. If drugs of this type are to be used to produce surfacing behavior of carp and possibly other fish, the less toxic the drug to man, the simpler would be the application of the method to field studies and to survival kits. Table 1 illustrates the order of activity of certain LSD-25 derivatives and congeners obtained on the nonpsychotic test group used for the past 5 years at this Laboratory. Note in Table 1 that LAE-32 is 1/30 as active in man as LSD-25. Other things being equal, it would appear that LAE-32 might well take the place of LSD-25 in experiments and field trials proposed.

C. Discussion

The mechanism of action of LSD-25 on the fish is unknown. As we have reported previously (4), certain oxidase poisons like sodium azide act

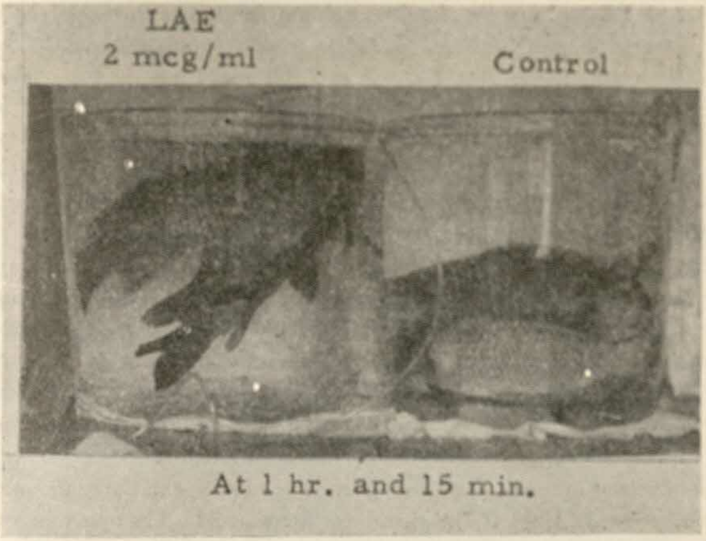


FIGURE 5

Illustrates that 2 mcg per ml of LAE-32 also affect the large fish, quantitatively speaking, the same way as LSD-25. There is a difference, however, in the rate at which LAE-32 reacts. LAE-32 takes longer to act and the larger carp do not remain as consistently at the surface as they do with a smaller concentration of LSD-25.

TABLE 1

ORDER OF ACTION OF LYSERGIC ACID DIETHYLAMIDE (LSD-25) DERIVATIVES IN MAN

		Data From Cold Spring Harbor
LSD-25	(d-lysergic acid diethylamide)	100
ALD-52	(1-acetyl d-lysergic acid diethylamide)	91
MLD-41	(1-methyl d-lysergic acid diethylamide)	36
LSM-775	(d-lysergic acid morpholide)	11
DAM-57	(d-lysergic acid dimethylamide)	11
BOL-148	(2-brom-d-lysergic acid diethylamide)	7
LPD-824	(d-lysergic acid pyrrolidide)	5
LAE-32	(d-lysergic acid ethylamide)	3

similarly to LSD-25 on the Siamese fighting fish. The surfacing behavior of the Siamese fighting fish may also be readily brought about by simple anoxia. It would appear that LSD-25 probably acts by poisoning one of the many oxidase systems which are found in living tissue, because the action of LSD-25 occurs whether the LSD-25 is injected intraperitoneally or placed in the outside liquid. As a matter of fact, the regression lines for the reaction-dose curves of LSD-25, whether in the outside liquid or injected, and Psilocybin when injected, are parallel.

Psilocybin has a much simpler chemical formula. The fact that injected Psilocybin acts like LSD-25 on the Siamese fighting fish (3) brings up the possibility that other compounds simpler than LSD-25 may be effective in inducing surfacing behavior. At present, the cost of LSD-25 prevents its use in large scale experiments. However, it is understood that a fermentation process used in Italy for producing d-lysergic acid might lead to a considerable reduction in the cost of the diethylamide derivative.

This question arises: Can fish, surfacing after exposure to minimal surfacing doses of LSD-25, be safely eaten by man? The question can be solved only by experimentation. However, the half-life of LSD-25 in warm-blooded animals is very short. Probably keeping the fish alive an hour or so after surfacing might lead to an almost complete disappearance of LSD-25 in the tissues. In this connection, experiments in this Laboratory concerned with boiling LSD-25 show that it is not rapidly destroyed by the heat of boiling water. However, we have no data on the effect of boiling LSD-25 in the presence of fish tissue. There is no doubt that LSD-25 does become inactive in the presence of light and of chlorine. The chlorine usually found in drinking water did not appreciably destroy LSD-25 in our earlier experiments (2).

In the preliminary experiments reported here, emphasis must be placed on the use of lysergic acid ethylamide (LAE-32) which is only 1/30th as active in man as LSD-25 (2). Rather large quantities of LAE-32 can be ingested by man without a psychosis developing. It would be of considerable importance to study the effects of LSD-25, LAE-32, and similar compounds in small ponds to determine if these materials or similar materials can be used as fish bait or pellets in survival kits. For example, a suitable pellet of LAE-32 dropped to the bottom of a lake inlet might provide sufficient edible fish, which might otherwise be unobtainable, by the surfacing phenomena.

Of importance to the problem of conservation is the feeding to fish of substances which would induce surfacing behavior. Experiments are planned in which LSD-25 and LAE-32 will be fed to carp to determine if absorption of these compounds takes place through the intestinal tract. Methods of forced feeding in carp have been worked out by Loeb and his co-workers (8).

D. SUMMARY

Although 1,800 assorted compounds force-fed to carp have not produced directed movements of the fish, the psychotomimetic compound d-lysergic acid diethylamide (LSD-25) as well as other derivatives and congeners do produce surfacing behavior. Preliminary data were obtained by exposing the fish to these compounds in the outside liquid. Carp, as much as 2 pounds in weight, responded, in general, to the same concentration of LSD-25 as did Siamese fighting fish weighing only one gram. Laboratory data are presented on goldfish, carp, and trout. It is proposed that extension of the present experiments to the field might lead to new methods of game conservation and the development of new formulations for survival kits.

REFERENCES

1. ABRAMSON, H. A., & EVANS, L. T. Lysergic acid diethylamide (LSD-25): II. Psychological effects on the Siamese fighting fish. *Science,* 1954, **120**, 990-991.
2. ABRAMSON, H. A. (*Ed.*). The Use of LSD in Psychotherapy. Transactions of a Conference on d-Lysergic Acid Diethylamide (LSD-25). New York: The Josiah Macy, Jr. Foundation, 1960.
3. ABRAMSON, H. A., ROLO, A., SKLAROFSKY, B., & STACHE, J. Production of cross-tolerance to psychosis-producing doses of lysergic acid diethylamide and Psilocybin. *J. of Psychol.,* 1960, **49**, 151-154.
4. ABRAMSON, H. A., WEISS, B., & BARON, M. O. Comparison of effect of lysergic acid diethylamide with potassium cyanide and other respiratory inhibitors on the Siamese fighting fish. *Nature,* 1958, **181**, 1136-1137.
5. EVANS, L. V., GERONIMUS, L. H., KORNETSKY, C., & ABRAMSON, H. A. Effect of ergot drugs on betta splendens. *Science,* 1956, **123**, 26.
6. LOEB, H. A. An electrical surface device for carp control and fish collection in lakes. *New York Fish and Game Journal,* 1955, **2**, 220-231.
7. ———. Night collection of fish with electricity. *New York Fish and Game Journal,* 1957, **4**, 109-118.
8. ———. Personal communication.
9. LOEB, H. A., & KELLY, W. H. A method for introducing capsules into the alimentary tract of carp. *New York Fish and Game Journal,* 1960, **7**, 160.

Long Island Biological Association
Cold Spring Harbor, New York

ADDENDUM: Since this paper was submitted for publication, the effect of feeding Psilocybin and Psilocin to goldfish has been observed. Both of these compounds are apparently absorbed through the intestinal tract of the fish when the drugs were incorporated in their food. Rather typical LSD effects are observed.

Published as a separate and in *The Journal of Psychology*, 1961, 52, 457-477.

THE ROLE OF PCF IN TEMPORAL MANIPULATIONS OF COLOR*[1]

Department of Psychology, Michigan State University

THOMAS M. NELSON AND S. HOWARD BARTLEY

A. INTRODUCTION

Bartley and Nelson (5) recently described color phenomena they designate by the terms "washout effect" or "color bleaching." *These terms describe the desaturation and large shifts in the hue of intermittent spectral targets at various PCF's and below CFF when these targets are compared with the same ones in the steady state.*

This phenomenon is not entirely new. Stewart reported the occurrence of hue changes with sub-fusional pulse rates in 1888 (11) and remarked at length upon the generality of such color phenomena in ordinary settings in 1924 (12). Walter, et al. (13, 14) also have more recently referred to hue changes. However, since washout effect is so sensitive to PCF, as we shall show, it is unlikely that the hue shifts produced were very pronounced in either case. Neither of these authors remarked about brightness and saturation changes.

To produce washout effect, Bartley and Nelson (5) employed five fairly discrete ranges of transmitted photic energies at four levels of illumination separated by log-steps. Stimulus intermittency was produced by an episcotister fitted with a two-open-sector disk. Observations made both by the authors and naive *O*s all showed that immediately or shortly after introduction of stimulus intermittency, brightness, saturation, and hue were altered. As rate was further increased, the shifts became even more pronounced, finally tapering off and reversing.

This experiment concerns the relation of PCF to washout effect. In the previous study a single PCF of .125 was used. In the present experiment PCF is to be systematically varied.

Completely quantitative results are not to be reported. We hope to be able to offer these in the future. We believe that further elucidation of the phenomena of washout effect prior to the possibility of full quantitative

---

* Received in the Editorial Office on August 2, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The present study is a part of the work being done under Research Grant NSF-G-5821 from the National Science Foundation.

characterization of the phenomena is worth something in itself since it may broaden the general understanding of color vision.

B. Apparatus and Experimental Procedure

The present study substituted an electronic apparatus producing square-wave pulses for the episcotister previously used. An electronic circuit manufactured by Grason-Stradler was used to drive a Sylvania glow-modulator tube. Because the luminous element of this gas tube has a small area, a biconvex lens was mounted between the tube and the eye in a way permitting a proper maxwellian view and thus increasing the effective area. We used a lens stopped to 0.5 inches in diameter. The lens face was observed 15 inches from the eye, providing a target subtending about 2° of visual angle.

The intensity of the source without a spectral filter was 3240 c/ft^2 at the highest level (intensity 3). Two other levels (intensities 2 and 1) were reductions of 1 and 2 log units respectively. These intensity reductions were achieved with Wratten neutral density filters. Two Wratten filters transmitting fairly narrow ranges of wavelength placed back of the lens produced red-orange and blue-green colors, respectively. They also further reduced the photic intensities involved. Specifically, Wratten filter "71A," the red-orange, transmitted 0.90 per cent of the impinging photic energies with dominant wavelengths between 630-640 mμ, cutting off all below 610 mμ. Wratten filter "75," the blue-green, transmitted 1.01 per cent and passed wavelengths in the 460-540 mμ range with 490 mμ its dominant wavelength.

Nine PCF's were used. These were .02, .03, .12, .25, .50, .75, .88, .97, and .98. Rates of intermittency, and hence the production of bleaching, were controlled by the observer. Intermittency rates were changed by a knob on the Grason-Stadler instrument.

Two practiced observers, the authors, served as principal observers. The task was fourfold. One task was to report the transition between flicker and fusion. The method of limits was employed here as elsewhere. Two ascending and two descending readings were taken. These data were averaged arithmetically to establish CFF indexes. A second task involved observing hue. The *O* reported (1) when the first changes in hue appeared as he increased rate from less than 1 pulse per second, (2) when a hue change first appeared as rate decreased from CFF plus 5 pulses per second, and (3) the rate at which hue changes appeared maximum using the slowest rate (less than 1 pulse per second) as the standard. This same procedure was followed for the variables of saturation and brightness. On these third and fourth tasks the observer likewise indicated the occurrence of first change, maximum change, and final change. Again maximum change was determined using the slowest rate obtainable on the Grason-Stadler as the standard. After obtaining numerical data under any condition, *O* manipulated rate and reported the nature of the changes. These were recorded so that the experimenter obtained both numerical and phenomenal data for each condition.

One thousand and eighty observations were made in all, i.e., 9 (PCF's) × 3 (intensities) × 2 (spectral conditions) × 5 (tasks) and 4 (observations). Distribution of observations varied slightly from one *O* to the other. The procedure for distributing observations followed recommendations arising from an earlier empirical study of CFF methodology (10).

Obviously no trials involved a fixed standard of reference. All were made by noting *changes of more or less* as stimulus intermittency was varied.[2] Lack of a fixed standard of reference, such as a monochromator, introduces a complication. It is possible that color shifts associated with a given PCF at the reference condition (less than 1 pulse per second) are not the same throughout. This implies that results of any one PCF may not be strictly comparable to results from any other. Later in the paper, the results of observing filtered targets with rate fixed at less than 1 pulse per second and PCF itself manipulated directly will be considered.

C. RESULTS

Quantitative data are given in Tables 1, 3, 5, 7. Figures 1-4 present data for the highest intensity condition occurring in each of the respective tables. Sensory end-results are given in Tables 2, 4, 6, 8. For most purposes, these presentations are complementary and best considered together.

---

[2] Essentially the same procedure was followed using 8 naive *Os* and the episcotister set-up previously reported (5). All *Os* gave responses indicative of washout effect. The results are not to be reported for the following reasons. First, *Os* verbalized uncertainty concerning their understanding of the task and capabilities for making the required discriminations. Recognition of color components in a flickering target is admittedly difficult at best and this difficulty is increased by our procedure of making successive rather than simultaneous comparisons. Second, the self-evaluations seemed borne out in the data. The full range of phenomena was rarely reported upon by an *O*. Most restricted themselves to reporting upon changes in hue and flicker. Naive *Os* have been found to be generally unsuited for sensory experiments. An experimental investigation (9) into the problem of differences between populations of observers leads one to expect the data of naive *Os* to show greater variability than those of practiced *Os* even though possessing the same central tendency. The data from 8 naive *Os* in this experiment in general confirmed this expectation.

TABLE 1

Stimulus Rates Producing Hue-Saturation, Brightness, Flicker-Fusion Changes for Nelson, Wratten 71A (Red-Orange)

		Intensity 3				Intensity 2				Intensity 1			
	PCF	Maximum	First	Terminal	CFF	Maximum	First	Terminal	CFF	Maximum	First	Terminal	CFF
Changes in brightness	.02	1.0	1.0	37.5	37.5	1.0	1.0	27.5	27.5	1.0	1.0	19.0	19.0
	.03	1.0	1.0	36.5	36.5	1.0	1.0	26.5	26.5	1.0	1.0	17.0	17.0
	.12	4.8	1.0	40.0	40.0	7.0	1.0	28.5	28.5	2.0	1.0	18.0	18.0
	.25	6.5	1.0	43.0	43.0	7.0	1.0	34.5	34.5	2.3	1.0	20.5	20.5
	.50	6.5	1.0	41.0	41.0	7.0	1.0	35.0	35.0	3.0	1.0	20.5	20.0
	.75	—	NBC[a]	—	40.0	—	NBC	—	33.0	—	NBC	—	19.0
	.88	—	NBC	—	33.5	—	NBC	—	29.0	—	NBC	—	15.0
	.97	—	NBC	—	18.0	—	NBC	—	16.0	—	NBC	—	9.0
	.98	—	NBC	—	16.0	—	NBC	—	13.0	—	NBC	—	7.0
Shifts in hue-saturation	.02	37.5	1.0	37.5	37.5	27.5	1.0	27.5	27.5	1.0	1.0	19.0	19.0
	.03	1.0	1.0	36.5	36.5	1.0	1.0	26.5	26.5	1.0	1.0	17.0	17.0
	.12	5.3	1.0	40.0	40.0	5.8	1.0	28.5	28.5	1.5	1.0	12.5	18.0
	.25	5.0	1.0	43.0	43.0	9.0	1.0	34.5	34.5	2.0	1.0	20.5	20.5
	.50	7.0	1.0	41.0	41.0	8.0	1.0	35.0	35.0	2.5	1.0	20.0	20.0
	.75	5.5	1.0	40.0	40.0	—	NHSS	—	33.0	—	NHSS	—	19.0
	.88	—	NHSS[b]	—	33.5	—	NHSS	—	29.0	—	NHSS	—	15.0
	.97	—	NHSS	—	18.0	—	NHSS	—	16.0	—	NHSS	—	9.0
	.98	—	NHSS	—	16.0	—	NHSS	—	13.0	—	NHSS	—	7.0

[a] NBC = No brightness change.

[b] NHSS = No hue-saturation shifts.

TABLE 2

Description of Hue-Saturation Shifts and Their Relation to Stimulus Frequencies and Intensities Reported for Wratten 71A (Red-Orange at Less than 1 Pulse Per Second) by Observer Nelson (See Table 1)

	Description of hue-saturation at indicated pulses per second		
PCF	Intensity 3	Intensity 2	Intensity 1
.02	orange (1.0), reddens (13.0), good red (37.5)	reddish-yellow (1.0), red (27.5)	gray-orange (1.0), gray (19.0)
.03	yellow (1.0), reddens (10.0), yellow absent, orange appears (13.0), orange-red (36.5)	white-yellow (1.0), orange (26.5)	gray-orange (1.0), gray (17.0)
.12	yellow-orange (1.0), red field halo (2.0), sun yellow (5.3), orange (10.5), orange deepens (30.3), clear orange (40.0)	orange-yellow (1.0), yellows (1.5), white-yellow (4.8), oranges (8.0), orange (29.5)	orange (1.0), whiter orange (1.5), orange (2.5)
.25	orange; red field halo (1.0), yellow; red field halo (2.0), sun yellow; red rim (5.0), clear yellow (11.0), oranges (12.0), clear orange (43.0)	orange (1.0), yellows (2.0), white-yellow (9.0), oranges (12.0), orange (34.5)	yellow-orange (1.0), yellow (2.0), orange (20.5)
.50	orange; magenta field halo (1.0), yellows (4.0), sun yellow (7.0), oranges (14.0), clear orange (41.0)	orange (1.0), yellows (4.0), yellow-orange (8.0), oranges (10.0) orange (35.0)	yellow-orange (1.0), yellow (1.5), yellow-pink (2.5), reddens (4.5), red-orange (20.0)
.75	orange (1.0), yellow orange (5.5), pale orange (40.0)	NHSS; (at 8.0 bright after image; light interval orange)	NHSS
.88	NHSS	NHSS	NHSS
.97	NHSS	NHSS	NHSS
.98	NHSS	NHSS	NHSS

TABLE 3
STIMULUS RATES PRODUCING HUE-SATURATION, BRIGHTNESS, FLICKER-FUSION CHANGES FOR BARTLEY, WRATTEN 71A (RED-ORANGE)

		Intensity 3				Intensity 2				Intensity 1			
	PCF	Max-imum	First	Ter-minal	CFF	Max-imum	First	Ter-minal	CFF	Max-imum	First	Ter-minal	CFF
Changes in brightness	.02	1.0	1.0	29.5	29.5	1.0	1.0	27.5	27.5	1.0	1.0	16.0	16.0
	.03	1.0	1.0	28.5	28.5	1.0	1.0	26.0	26.0	1.0	1.0	20.0	20.0
	.12	1.0	1.0	37.5	37.5	1.0	1.0	27.0	27.0	1.0	1.0	18.5	18.5
	.25	1.0	1.0	37.0	37.0	2.5	1.0	31.0	31.0	1.0	1.0	19.0	19.0
	.50	—	NBC	—	37.0	—	NBC	—	33.5	—	NBC	—	21.0
	.75	—	NBC	—	36.5	—	NBC	—	34.0	—	NBC	—	19.0
	.88	—	NBC	—	32.0	—	NBC	—	27.0	—	NBC	—	15.5
	.97	—	NBC	—	14.5	—	NBC	—	14.0	—	NBC	—	9.0
	.98	—	NBC	—	14.5	—	NBC	—	9.0	—	NBC	—	6.0
Shifts in hue-saturation	.02	29.5	1.0	29.5	29.5	10.0	3.6	10.0	27.5	16.0	1.0	16.0	16.0
	.03	28.5	1.0	28.5	28.5	11.3	2.0	11.3	26.0	20.0	1.0	20.0	20.0
	.12	5.0	1.0	32.0	27.5	17.0	1.0	17.0	27.0	18.5	1.0	18.5	18.5
	.25	15.0	1.0	23.0	37.0	1.5	1.0	18.0	31.0	19.0	1.0	19.0	19.0
	.50	8.0	1.0	37.0	37.0	5.0	1.0	33.5	33.5	—	NHSS	—	21.0
	.75	—	NHSS	—	36.5	5.0	1.0	5.0	34.0	—	NHSS	—	19.0
	.88	—	NHSS	—	32.0	—	NHSS	—	27.0	—	NHSS	—	15.5
	.97	—	NHSS	—	14.5	—	NHSS	—	14.0	—	NHSS	—	9.0
	.98	—	NHSS	—	14.5	—	NHSS	—	9.0	—	NHSS	—	6.0

TABLE 4

Description of Hue-Saturation Shifts and Their Relation to Stimulus Frequencies and Intensities Reported for Wratten 71A (Red-Orange at Less Than 1 Pulse Per Second) by Observer Bartley (See Table 3)

	Description of hue-saturation at indicated pulses per second		
PCF	Intensity 3	Intensity 2	Intensity 1
.02	orange-red (1.0), reddens (2.5), clear red (29.5)	orange-red (1.2), dark red (15.0)	orange (1.0), gray (16.0)
.03	yellow-red (1.0), orange-red (2.0), red (28.5)	red (1.0), red-orange (3.5), yellow-orange (6.0), darker orange-red (10.0)	orange (1.0), gray (20.0)
.12	yellow-orange, gamma movement, third dimension movement (1.0), sun-yellow with thin red halo (5.0), gray yellow-orange (14.0), clear orange (32.0)	orange (1.0), reddens (17.0)	orange (1.0), lighter orange (2.0), gray-orange (18.5)
.25	orange with red field halo (1.0), yellow-orange, thin red halo (5.0), sun-yellow (15.0), clear orange (23.0)	orange (1.0), pale orange (1.5), clear orange (18.0)	orange (1.0), gray-orange (19.0)
.50	orange (1.0), yellow (8.0), oranges (20.0), clear orange (37.0)	orange (1.0), yellow-orange (5.0), orange (33.5)	NHSS
.75	NHSS	orange (1.0), pale orange (5.0)	NHSS
.88	NHSS	NHSS	NHSS
.97	NHSS	NHSS	NHSS
.98	NHSS	NHSS	NHSS

TABLE 5

STIMULUS RATES PRODUCING HUE-SATURATION, BRIGHTNESS, FLICKER-FUSION CHANGES FOR NELSON, WRATTEN 75 (BLUE-GREEN)

		Intensity 3				Intensity 2				Intensity 1			
	PCF	Maximum	First	Terminal	CFF	Maximum	First	Terminal	CFF	Maximum	First	Terminal	CFF
Changes in brightness	.02	8.0	1.0	29.5	29.5	7.0	7.0	20.0	20.0	1.0	1.0	5.0	12.0
	.03	11.0	1.0	25.5	25.5	1.0	1.0	16.5	16.5	1.0	1.0	4.5	11.0
	.12	2.0	1.0	34.0	34.0	7.0	1.0	16.0	16.0	1.0	1.0	5.0	11.5
	.25	2.5	1.0	32.0	32.0	1.0	1.0	7.0	27.5	1.0	1.0	7.0	14.0
	.50	2.3	1.0	37.0	37.0	1.0	1.0	1.5	26.0	1.0	1.0	7.0	14.0
	.75	—	NBC	—	37.0	1.0	1.0	8.0	25.5	1.0	1.0	8.0	13.5
	.88	—	NBC	—	31.0	—	NBC	—	22.5	—	NBC	—	11.0
	.97	—	NBC	—	20.0	—	NBC	—	12.5	—	NBC	—	7.0
	.98	—	NBC	—	17.0	—	NBC	—	10.5	—	NBC	—	4.5
Shifts in hue-saturation	.02	29.5	1.0	29.5	29.5	20.0	1.0	20.0	20.0	4.0	1.0	4.0	12.0
	.03	25.5	1.0	25.5	25.5	3.0	1.0	16.5	16.5	4.5	1.0	4.5	11.0
	.12	2.0	1.0	34.0	34.0	7.0	1.0	16.0	16.0	5.0	1.0	5.0	11.5
	.25	2.0	1.0	32.0	32.0	4.0	1.0	7.0	27.5	—	NHSS	—	14.0
	.50	2.3	1.0	37.0	37.0	1.0+	1.0	1.5	26.0	—	NHSS	—	14.0
	.75	—	NHSS	—	37.0	—	NHSS	—	25.5	—	NHSS	—	13.5
	.88	—	NHSS	—	31.0	—	NHSS	—	22.5	—	NHSS	—	11.0
	.97	—	NHSS	—	20.0	—	NHSS	—	12.5	—	NHSS	—	7.0
	.98	—	NHSS	—	17.0	—	NHSS	—	10.5	—	NHSS	—	4.5

TABLE 6

Description of Hue-Saturation Shifts and Their Relation to Stimulus Frequencies and Intensities Reported for Wratten 75 (Blue-Green at Less Than 1 Pulse Per Second) by Observer Nelson (See Table 5)

	* Description of hue-saturation at indicated pulses per second		
PCF	Intensity 3	Intensity 2	Intensity 1
.02	blue (1.0), gray (29.5)	gray-blue (1.0), gray (20)	gray-blue (1.0), gray (4.0)
.03	blue (1.0), gray-blue (25.5)	blue (1.0), lighter blue (3.0), gray-blue (16.5)	gray-blue (1.0), gray (4.5)
.12	blue (1.0), gray-white, red tinge (2.0), blue (34.0)	blue (1.0), lighter blue (7.0), blue (16.0)	blue (1.0), dark blue (5.0)
.25	blue (1.0), white-blue (2.0), blue (32.0)	blue (1.0), lighter blue (4.0), blue (7.0)	NHSS
.50	blue (1.0), white-blue (2.3), blue (37.0).	blue (1.0), lighter blue (1.0+), blue (1.5)	NHSS
.75	NHSS	NHSS	NHSS
.88	NHSS	NHSS	NHSS
.97	NHSS	NHSS	NHSS
.98	NHSS	NHSS	NHSS

TABLE 7

STIMULUS RATES PRODUCING HUE-SATURATION, BRIGHTNESS, FLICKER-FUSION CHANGES FOR BARTLEY, WRATTEN 75 (BLUE-GREEN)

		Intensity 3				Intensity 2				Intensity 1			
	PCF	Max-imum	First	Ter-minal	CFF	Max-imum	First	Ter-minal	CFF	Max-imum	First	Ter-minal	CFF
Changes in brightness	.02	1.0	1.0	15.5	15.5	1.0	1.0	13.0	13.0				
	.03	1.0	1.0	21.5	21.5	5.0	5.0	12.5	12.5				
	.12	3.3	1.0	28.0	28.0	2.6	2.6	17.5	17.5	1.0	1.0	6.0	6.0
	.25	8.0	1.0	30.0	30.0	2.9	2.8	8.0	20.5	2.0	2.0	9.0	9.0
	.50	4.0	4.0	29.0	29.0	5.4	8.3	13.0	22.8	8.0	8.0	12.0	12.0
	.75	1.0	1.0	30.5	30.5	4.8	3.9	13.2	24.7	5.0	5.0	10.0	10.0
	.88	—	NBC	—	25.0	5.8	5.1	13.7	18.5	3.5	3.5	8.0	8.0
	.97	—	NBC	—	18.0	—	NBC	—	11.0	5.0	5.0	5.0	3.5
	.98	—	NBC	—	16.0	—	NBC	—	8.5	—	NBC	—	3.0
Shifts in hue-saturation	.02	15.5	1.0	15.5	15.5	13.0	1.0	13.0	13.0				
	.03	1.0	1.0	21.5	21.5	12.5	1.0	12.5	12.5				
	.12	5.0	1.0	23.0	28.0	3.3	3.3	13.0	17.5	6.0	1.0	6.0	6.0
	.25	3.0	1.0	14.0	30.0	2.6	1.0	11.4	20.5	6.0	1.0	6.0	9.0
	.50	5.5	1.0	13.0	29.0	5.5	1.0	9.6	22.8	12.0	1.0	12.0	12.0
	.75	—	NHSS	—	30.5	2.4	1.0	5.6	24.7	5.0	1.0	10.0	10.0
	.88	—	NHSS	—	25.0	6.4	1.0	11.5	18.5	—	NHSS	—	8.0
	.97	—	NHSS	—	18.0	—	NHSS	—	11.0	—	NHSS	—	3.5
	.98	—	NHSS	—	16.0	—	NHSS	—	8.5	—	NHSS	—	3.0

TABLE 8

Description of Hue-Saturation Shifts and Their Relation to Stimulus Frequencies and Intensities Reported for Wratten 75 (Blue-Green at Less Than 1 Pulse Per Second) by Observer Bartley (See Table 7)

	Description of hue-saturation at indicated pulses per second		
PCF	Intensity 3	Intensity 2	Intensity 1
.02	blue (1.0), gray-blue (15.5)	dim blue (1.0), gray-black (13.0)	
.03	purple (1.0), blues (3.0), blue (21.5)	blue (1.0), dark gray (12.5)	
.12	red-purple (1.0), white (5.0), red-purple (6.5), less red (8.0), purple absent (12.0), green-blue (23.0)	blue (1.0), blue lightens (3.5), dark brown-blue (13.0)	green-gray (1.0), gray-black (6.0)
.25	blue (1.0), blue-white (3.0), purple-blue (7.0), blue (14.0)	blue-green (1.0), yellow-blue (3.0), blue (10.5)	green-blue (1.0), grays (3.0), gray (6.0)
.50	blue (1.0), purple-blue (5.5), blue (13.0)	green-blue (1.0), desaturating, yellow present (5.5), blue-green (9.0)	blue (1.0), gray-blue (12.0)
.75	NHSS	blue-green (1.0), blue fades, green slightly more prominent (3.5)	blue (1.0), light blue (5.0), gray-blue (10.0)
.88	NHSS	light blue-green (1.0), slight desaturation (5.0)	NHSS
.97	NHSS	NHSS	NHSS
.98	NHSS	NHSS	NHSS

Inspection of the tables shows that many of the conditions produced color changes. These seem to be of two sorts, however. Some are properly termed washout effect and others are reminiscent of changes producible by sheer manipulation of photic intensity. It is well to remember that washout effect denotes desaturation and shift in hue produced by manipulation of repetitive temporal distributions of photic energies. Sometimes the manipulation of rate increases saturation. Sometimes hue shifts toward gray as

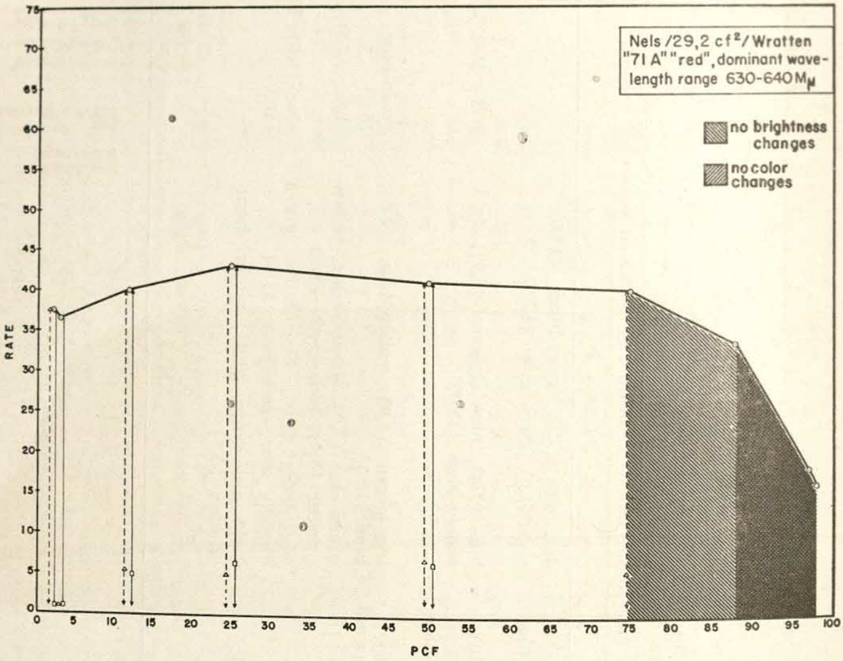


FIGURE 1

Range within which color changes occur. Circles indicate CFF level. Triangles show locale at which maximal shifts in hue-saturation occur. Squares show locale of maximal brightness. Dotted lines extending from the triangles and squares indicate the range of changes.

fusion is reached. Such color changes fall outside those we have designated as washout effects. Many of the color changes associated with PCF's .02 and .03 as well as virtually all of *B*'s and many of *N*'s responses under the condition of minimal intensity are examples. Since reducing PCF reduces the over-all energies exciting the retina, one may tentatively conclude that washout phenomena are critically dependent upon definite levels of intensity.

Other color effects seemed clearly to depend upon timing of the photic input. In fact, forty-three reports indicated washout effect. Tables 1, 3, 5, 7 give the region of maximum, first, and terminal washout effects as accurately as we were able to determine them without a matching source. In every case, desaturation and shift in hue became obvious as soon as rate was increased from our minimal rate, if they were to occur at all. In forty of the forty-three cases washout effect was initiated at 1 pulse per second.

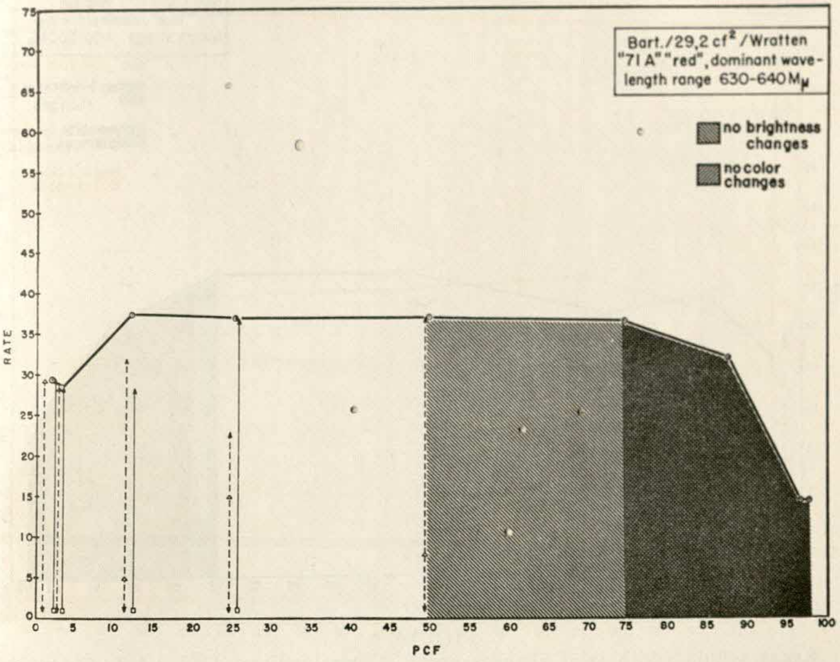


FIGURE 2

Range within which color changes occur. Circles indicate CFF level. Triangles show locale at which maximal shifts in hue-saturation occur. Squares show locale of maximal brightness. Dotted lines extending from the triangles and squares indicate the range of changes.

The terminal point of washout effect was variable. Generally the higher the CFF value, the more extensive the range of washout effect. Color seemed to stabilize when CFF was reached. See Figures 1 and 3. Maximum washout effect, with but a single exception, occurred with the first half of the subfusional intermittency range. Median and modal rates for maximum washout were 5.0 pulses per second and the mean rate was 4.8. According to the measures we used the maximum washout effect rose detectably as intensity declined.

The location of maximum washout effect did not seem affected by PCF. The various PCF's did vary greatly, however, in their effectiveness in producing washout effect. The phenomenon occurs most reliably with moderately small PCF's. Both *O*s consistently reported red washout effect for PCF's .12 and .25, and blue washout effect for PCF's .12, .25, and .50 at the upper intensity conditions (intensities 3 and 2). The frequency of reports indicating washout effect declined at PCF .50 for the red and .97 for the blue. At .97 and .98 color changes are never reported. Absence of reports should not necessarily be interpreted to mean that subfusional color changes were absolutely non-existent under these conditions. Slight changes could almost certainly slip through unnoticed by our method. The results do show

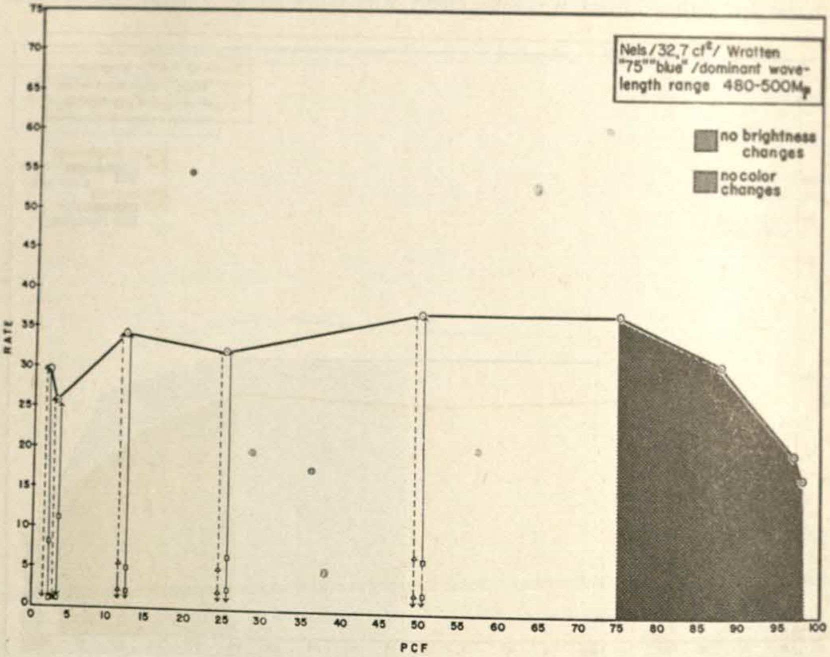


FIGURE 3
Range within which color changes occur. Circles indicate CFF level. Triangles show locale at which maximal shifts in hue-saturation occur. Squares show locale of maximal brightness. Dotted lines extending from the triangles and squares indicate the range of changes.

that changes in hue and saturation that were blatant with small PCF's, with large PCF's became at least so small that a matching source would be necessary for their detection.

When the figures and tables containing quantitative data are paired with their appropriate even-numbered tables a common sequence of events is detectable. If washout is to occur it begins at once. The immediate change is such that what is seen as red under steady transmittency becomes orange, and what is seen as blue-green becomes blue. Further increase toward maximum washout effect produces shifts toward yellow in the first case and white in the second. At maximum, the red becomes a glaring white-yellow disk and the blue-green, a white-blue disk. Alterations in saturation co-vary with those of hue. Desaturation was always seen as maximum in the region of maximum shift in hue. Since increasing intermittency rate beyond the

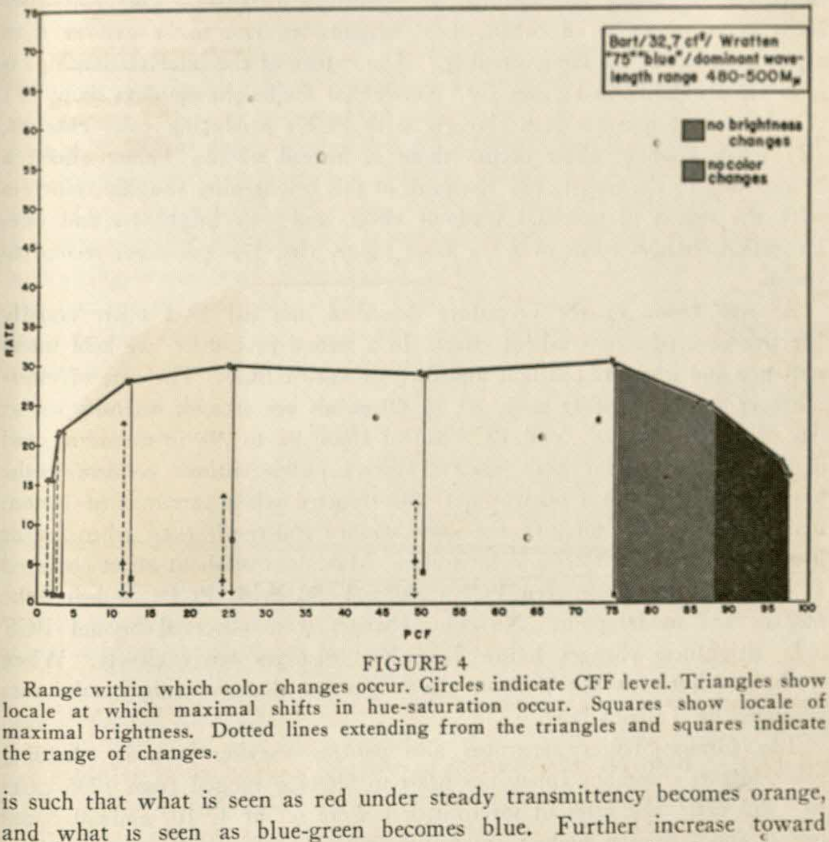


FIGURE 4

Range within which color changes occur. Circles indicate CFF level. Triangles show locale at which maximal shifts in hue-saturation occur. Squares show locale of maximal brightness. Dotted lines extending from the triangles and squares indicate the range of changes.

region of maximum washout effect reverses the sequence of sensory events, columns of data separating hue and saturation were not prepared.

In addition to changes located in the target itself, field halos (probably the effect of stray illumination on the retina) and bright highly saturated edges were at times reported.

Bartley and Nelson (5) previously reported that brightness changes are associated with washout effects, and this is shown here too. Hue-saturation shifts are frequently accompanied by brightness increases. These increases often appeared to be of enhancement magnitudes, i.e., to be greater than a steady target of the same intensity.[3] The nature of the interrelationships is most plainly shown in Figures 1-4. Notice that the brightness data show (1) the shifts in brightness occur largely with PCF's producing color changes, (2) when washout effect occurs there is, instead of the Talbot effect, a brightening of the target, (3) the peak of this brightening roughly coincides with the region of maximal washout effect, and (4) brightness and hue-saturation changes occur over the same range, viz., 1+ pulse per second to fusion.

As was found by the procedure described, not all PCF's are equally effective in producing washout effect. In a second procedure, we held intermittency and intensity constant and directly varied PCF. The rate of intermittency was held either at 1, 10, or 60 pulses per second, intensity set at one of our three levels, and PCF varied from .02 to .98 in ascending and descending orders for both spectral filters. Observations confirmed the previous results. Thus no washout was reported when rate was at fusion, and all PCF's did not give the same sensory end-result even when set at identical subfusional rates and intensities. Maximum washout effect occurred at highest intensity between PCF's .06 and .30, with PCF .24 being the median and modal point. No color changes were observed beyond PCF .64. Brightness changes followed washout changes fairly closely. When washout effect was not present, the brightness changes were of the sort predicted by Talbot's law.

The Grason-Stadler apparatus also permits varying intensity, holding other factors constant. Intensities prior to filtering ranged from 1296 c/ft^2 to 5962 c/ft^2 and rates of intermittency were set at 1, 10, and 60 cycles per second. At the highest intensities, washout effect of red seemed to progress toward white. Other than this, intensity had only the general effect of desaturating the blue and shifting red toward yellow. This is what one might expect on the basis of the Bezold-Brücke phenomenon. The shifts were most pronounced for the large PCF's since greater total luminance per unit of time was involved. The smaller PCF's did not give so large a range of color phenomena and dull slightly colored grays predominated. Washout effect is therefore to be distinguished from color phenomena produced by intensity shifts.

---

[3] Our equipment at that time did not permit certain identification of enhancement. This possibility is presently being tested by R. Ball using a monochromatic matching field but it is too early to speculate upon the outcome of this test.

D. DISCUSSION

The general inference to be made from the over-all results is that timing of stimulus input is a critical variable in producing washout effect, and thus, by implication, critical also in producing the sensory end-result, color. While it is customary to treat color as independent of temporal factors in impingement, there is no necessary reason for doing so. Custom alone has seemed to dictate this and consequently minimize the manipulation of temporal factors such as duration, repetition frequency of brief pulses, PCF, etc. Thus convention has restricted color theories to phenomena produced by impingements of unlimited duration. Surely whatever the nature of underlying events eventuating in a sensory end-result of color, they require time to run to completion. The study of the effects of various types of impingement timing is as important to our understandings as investigations of spatial variables or retinal anatomy.

At that, the general fact is perhaps not so surprising as the specific. While brightness, hue, and saturation might be expected to vary as a function of pulse duration, the details of the washout process seem to involve something other than mere threshold. In intermittent stimulation, some ratios of pulse duration to intervals between them (PCF's) effectively bleached color and others were never observed to do so under any condition. Under the conditions used, a wide range of intermittencies and a wide range of PCF's were involved. Within this combination of factors, durations of given values were variously coupled with the factors just mentioned. Some cases produced washout effect, and some did not, hence it would seem that pulse duration per se was not the crucial factor. If duration per se were the crucial variable, and if a PCF .25 produced washout effect at 5 cycles per second, as it did, then a PCF of .96 should do so at about 19 cycles per second since at that rate pulse durations would be equal. This did not of course happen. PCF .96 produced washout effect nowhere for either *O* at any rate between 1 and 100 pulses per second. Washout effect seems to depend upon successive stimulation and therefore upon whatever cumulative effects are produced in the visual system as a result of repetitive activation of nervous units. This being the case, let us consider the critical effects produced in the visual system by cyclic forms of stimulation. These may be related to the washout effect by analogy.

Notice that the indexes of CFF given in Tables 1, 3, 5, 7 do not decline uniformly as PCF is increased, although this surely might well be the naive expectation. As PCF becomes greater, the period of no stimulation becomes progressively shortened and presumably easier to bridge with continuous activity. Therefore the traditional explanation has been that, all other variables being equal, as PCF is increased CFF should decline. This does not happen. CFF increases or does not change. Coupled with this seeming paradox, the washout effect is associated almost solely, and irrespective of photic wavelength characteristics, with those portions of the curve that do not decrease as expected.

Now in spite of the paradoxical nature of the CFF increase, such a finding is not entirely new (1). Indeed curves comparable in shape to those described by the present data have recently been reported (6, 7, 8) for full spectrum stimulation under roughly comparable conditions of intensity, area, etc. This gains added significance since a theoretical explanation has been given for the form the full spectrum curves assume.

Thus, in resumé, we have (1) several sets of curves presenting an unexpected form, (2) washout effects occurring only with the paradoxical portions of the curves, and (3) a theoretical explanation for such outcomes based upon the results obtained under full spectrum conditions.

These results lead us to consider essentially what happens in any case when the sensory outcome changes from flicker to steady brightness (fusion). This would form a good basis for understanding all features of sensory outcomes that depend upon temporal manipulations of impingement.

Fusion is manifested at the sensory level as a continuous brightness. The neurophysiological correlate is some uniform rate of cortical discharge occurring at some threshold of density. Conversely, all that is needed to produce flicker is a temporally irregular discharge of elements. Therefore reduction of PCF while holding other factors constant will progressively reduce CFF, unless, in addition to elements already in operation, sizeable populations of new elements having very different latencies, or transmission rates, or inhibitory functions come into operation over a part of the cycle. That such is the case is well known. Much is said in the literature nowadays about on- and off-response elements, but the conditions for their appearance, and the nature of their interactions have not fully been worked out.

The previous article discussed the influence various rates of transmission might have upon washout effects as various samples of the spectrum were compared. In this discussion it seems most appropriate to center on latency and the conditions under which both on- and off-discharges can be inhibited, for it is possible that PCF is dependent upon such factors.

Evidence from over-all recordings shows on-elements require less total energy to effect discharge than off-elements. In addition, cortical off-discharges do not seem to increase the inhibition of on-discharges when the two are mixed. Bartley (3) reports that a cortical on-discharge can follow a cortical off-discharge as closely as 12 msec., whereas one on-discharge cannot follow another unless they are at least 80 msec. apart. These facts suggest that suitably long stimulus durations introject off-discharges into what would otherwise be a constant level of discharge. As a result flicker is reinstituted. This fact was used to account for the unexpected rise in CFF as PCF was lengthened (8).

Decline in CFF with increase in PCF finally occurs. This can be attributed partly to inhibition of off-discharge. When suitable durations and intensities producing off-discharges are reached, these discharges can be obliterated by shortening the intervals between pulses by increasing PCF (6, 7, 8). Cortical off-discharge is further linked to washout effect by observations at the physiological level. Bartley (3) reports that very definite on- and off-discharges occur when a slow photic pulse rate is coupled with a PCF of .25. This is what was reported in our previous article as being virtually the optimal PCF for washout effect.

If cortical off-responses do play a role both in washout effect and extending the range of flicker as has been suggested, a puzzle emerges. The very mechanism allowing for eminently fine detection of photic intermittency plays a role in the coarsening of wavelength discrimination.

Washout effect was also found related to target brightening. Both occurred at approximately the same ranges of photic intermittency and with approximately the same range of PCF's. Other researchers have also reported brightening within this same range of intermittencies. Therefore washout effects may be associated with the completeness of utilization of channels within the optic pathway, as well as with the propagation of cortical off-discharges. This involves one with the *alternation of response theory* (2, 4) and hence more detail than can be dealt with here. Simply let it be said that there is a fixed number of parallel channels in the optic pathway[4] available to transmit retinal discharges and that each optic nerve fiber is the initial section of one of these channels. The nature of the central end of a channel is not nearly so obvious. It is known from experimentation that one is periodically able to activate a greater proportion of the finite number of central channels by timing of photic stimuli and that when intensities are at a proper level this is seen as brightening. An alternative means of bringing about more complete utilization of central channels and thence increasing brightness is simply to increase stimulus intensity. Either means will apparently worsen color discrimination. The first produces washout effects and the second the Bezold-Brücke phenomenon. Therefore one might suppose that complete utilization of the central pathway tends to produce achromatic vision. It would be interesting in this regard to see whether a "washedout" target can be matched simply be increasing the intensity of a steady source transmitting the same selected wavelengths as the one bleached temporally. If our supposition is true, it should.

---

[4] The optic pathway consists of: (a) the tract, including the optic nerve; (b) the relay nuclei of the lateral geniculate, the pretectal area and the superior colliculus; (c) the radiation to cortex, and paths to thalamus and tectal area; and (d) the projection areas including the cortical, thalamic and tectal projections. In addition to these, there are (e) the association areas. Just where to delimit the visual system is problematic, depending upon how one views neural functions.

E. Summary

The sensory end-result called "washout effect" was related to various temporal features describing intermittent stimulation. Pulse-to-cycle fraction was found to be a critical variable in producing a washout effect at either end of the normally colored spectrum. The experimental outcomes suggest that washout effect is associated with propagation of cortical off-discharge and full utilization of the optic tract.

References

1. Bartley, S. H. Vision: A study of its basis. New York: Van Nostrand, 1941.
2. ———. Some facts and concepts regarding the neurophysiology of the optic pathway. *A.M.A. Arch. Ophth.*, 1958, **60**, 775-791.
3. ———. Central mechanisms in vision. *In* Field, J., Magoun, H. W., & Hall, V. G., (*Eds.*), *Handbook of Physiology, Sec 1, Neurophysiology, Vol. I.* Washington, D.C.: *Amer. Physiol. Soc.*, 1959. (Chap. 30, Pp. 713-741.)
4. ———. Clarification of some of the procedures and concepts involved in dealing with the optic pathway. *In* Jung, R., (*Ed.*), *Visuelles Symposium.* Heidelberg: Springer-Verlag, 1961.
5. Bartley, S. H., & Nelson, T. M. Certain chromatic and brightness changes associated with rate of intermittency of photic stimulation. *J. Psychol.*, 1960, **50**, 323-332.
6. ———. Equivalence of various pulse-to-cycle fractions producing critical flicker frequency. *J. Opt. Soc. Amer.*, 1960, **50**, 241-244.
7. ———. Some relations between pulse-to-cycle fraction and critical flicker frequency. *Percept. Mot. Skills*, 1960, **10**, 3-8.
8. ———. A further study of pulse-to-cycle fraction and critical flicker frequency. A decisive test. *J. Opt. Soc. Amer.*, 1961, **51**, 41-45.

9. NELSON, T. M., BARTLEY, S. H., & DEHARDT, D. A comparison of variability of three sorts of observers in a sensory experiment. *J. Psychol.*, 1960, **49**, 3-11.
10. ———. Considerations involved in obtaining averages of CFF data. *J. Psychol.*, 1960, **49**, 263-277.
11. STEWART, G. N. Is the law of Talbot true for every rapidly intermittent light? *Proc. Roy. Soc. Edinb.*, 1887-1888, **15**, 441-464.
12. ———. Color phenomena caused by intermittent stimulation with white light. *Amer. J. Physiol.*, 1924, **69**, 337-353.
13. WALTER, V. J., & WALTER, W. G. The central effects of rhythmic sensory stimulation. *EEG Clin. Neurophysiol.*, 1949, **1**, 57-86.
14. WALTER, W. G., DOVEY, V. J., & SHIPTON, H. Analysis of the electrical response of the human cortex to photic stimulation. *Nature*, 1946, **158**, 540-541.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 479-490.

THE EFFECT OF AREAL CHARACTERISTICS OF TARGETS UPON SHAPE-SLANT INVARIANCE*

Department of Psychology, Michigan State University

THOMAS M. NELSON, S. HOWARD BARTLEY, AND CHARLES M. BOURASSA

A. PROBLEM

There is good reason to suspect that what is seen as shape is influenced by the character of internal areas. Orbison (23, 24) showed that extension of various kinds of field structure into the area bounded by an outline can affect shape. One of the questions to be posed here concerns the effect that surfaces bounded by a continuous edge have upon shape. One might expect shape to vary as a function of the degree of differentiation within the bounded area.

Assessing the effect of such a variable upon shape[1] for theoretical purposes necessarily involves considering what is seen as slant. It is therefore appropriate to digress enough to give the two most prominent reasons slant must be considered.

When working with responses having a known physical analogue, one tends to use these physical analogues in theoretical description and explanation. The nature of the stimulating situation enters one's explanation of the perceptual responses elicited. The shape constancy indexes suggested by Brunswik (4) and Thouless (27) are examples. Targets to which the *O* responds are described in terms of "distal" ("real") and "proximal" ("stimulus") properties. The "distal" or "real" shape (form) is a physical[2] description of the target from a point of reference 90° to the surface. "Prox-

---

* Received in the Editorial Office on August 2, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] *Shape* refers to what is seen; *form* refers to the target as a physical area. *Slant* refers to the observed orientation of the object; and *inclination*, to the relation of the target to the perpendicular.

[2] "Physical" here may be taken to mean a Newtonian or Euclidean description of space. It is widely held that Euclidean geometry provides a satisfactory description of "space" of the magnitude we encounter on earth although it takes non-Euclidean geometry to describe "space" on a larger scale, i.e., "space of the universe." This is a convenient analogue and will do for the present. It is, however, fundamentally false. Geometry, no matter how qualified (i.e., as Euclidean, non-Euclidean, synthetic, analytic, etc.), is an *abstract mathematical discipline interpretative of space and not a descriptive generality having an empirical (observational) character* (8, 16). The relation of geometrical interpretation to perceptual description will be reintroduced in a more sophisticated manner later in this paper.

imal" or "stimulus" shape (form) is a physical description of target using the *O* as the physical point of reference. Physical form necessarily involves a physical point of reference. It is impossible to describe one property without dealing with the other.

Perceived shape is linked with slant for the *O*. A flat shape is always perceived to be on some plane. The importance of slant perception for shape perception becomes particularly evident when one is concerned with effective overt response.

When discriminations of shape and discriminations of slant are related directly (19, 21) the relationship between the properties appears to be lawful. The relation is very different, however, from the geometric description of the target situation. Geometric invariances are generally poorly related to psychological invariances. Research has failed to demonstrate an invariant shape-slant relation (3, 6, 17, 18, 26).

Shape-slant invariance need not involve Euclidean space. Despite this, there are those who still hope to show an invariant relation between shape-slant responses and geometric function. From curves such as those of Figures 1e and f, however, one would predict the failures that have in fact ensued.

Koffka's (17) is the best-known assertion of the traditional expectation that geometry and sensory processes tally, and the best-known test of this is Stavrianos' (26). To date, all psychological approaches to analysis of shape-slant data and statements of shape-slant theory have been based upon some such assumption. Our presentation criticizes the idea that man's natural discriminations contain invariance relations similar to those ascribed to the world by common geometry. The axioms of ordinary geometry may not be taken to be perceptual or common sense necessities (Clifford, 8) (Helmholtz, 16).

Our aim in this study is to consider changes in shape and slant as sensory responses, not as geometrical items, and to show the invariant relation between them. We hope to show that the diverse natures of the areas within the target forms affect the *O*'s shape-slant discrimination and thereby the nature of the invariant relation.

B. MATERIALS, SUBJECTS, AND PROCEDURE

Luminous targets were presented to *O* in an unilluminated field. Two contiguous darkrooms were used; one for the *O* and the other for the exposure-stage and the presentation equipment. All surfaces of both rooms and the stage were painted "flat black." A window 14 inches high and 28 inches wide, connected the two rooms. Its center was at eye-level when the *O* was seated. The window was closable as needed, by a black draw-curtain. Distance from eye to target was 17-1/3 feet.

The targets, covered with luminescent paint, were made visible by an ultraviolet source. Four series of targets differing in areal features (seen as texture) were used. Each series contained 20 targets differing in inclination and geometric properties. The targets were 5 × 5 inch circles and 5 × 4, 5 × 3, 5 × 2 and 5 × 1 inch ellipses. All groups of circles and ellipses had 4 members each, with each member mounted in a different orientation with reference to the *O*. One of each form was mounted in the vertical (0°); one was tilted 22.5° from the vertical (away from the *O* at the top); one was at 45°, and one at 67.5°. One series of outline targets of 12 gauge wire, a second series of flat, untextured surfaces, a third series, of fine-textured surfaces, and a fourth, of coarse-textured surfaces produced the areal differences. Each target was viewed binocularly with naked eye through a 2-1/4 × 3-1/4 inch aperture placed 30 inches away.

Forty *O*s were used, 10 for each areal experiment. The task of each *O* was twofold. One experimental run required reproduction of the inclination of the target by a "tilt-board." The tilt-board (10 × 7-1/2 inches) was pivoted on the extended shaft of a Variac (autotransformer) whose output led to a voltmeter in the experimenter's room. The voltmeter was calibrated in units of inclination or tilt. The tilt-board was covered by a "white" matte card surface visible in the dim illumination of a 7-watt red bulb overhead. The instructions were as follows: "You are to be presented visually a series of forms or figures, some of which are mounted in positions other than upright. Your task is to duplicate the plane in which the various figures lie by manipulating the tilt-board in front of you until the plane at which you set the tilt-board best approximates the plane in which you perceive the object to be oriented."

In the other experimental run, the *O*s were called upon to draw the shapes of the targets. The instructions were as follows: "You are to be presented a series of visual forms or figures. Your task is to draw the form that you see from your perspective, being particularly careful to keep the minor/major axis proportionality as you see it."

*O*s were allowed unlimited time to make judgments and drawings, but these were usually made promptly. Every condition was distributed randomly throughout a given series. Half of the *O*s in each experiment started with drawing and half with slant indications.

C. Results and Discussion

The results of the experiments are given in Figures 1-3. The appropriate ones indicate a decrease in shape index (ratio of minor to major axis) as slant increases, clearly confirming the expectation that shape and slant change in a constant or invariant fashion. They demonstrate that shape-slant invariance is formed by the responses themselves in the absence of ordinary conceptual intermediaries, i.e., the "real" Euclidean conditions of stimulation.

For the outline targets the invariance is almost linear over the whole shape-slant continuum. This is readily seen when the points plotted in Figure 1a are noted in relation to the diagonal line representing true linearity. This diagonal takes into account the whole range of possibilities and obliterates the distinctions between targets as different classes of stimuli, i.e., as being 1 × 5, 2 × 5, 3 × 5, etc., and as being inclined to various extents. Nelson and Bartley (21) earlier reported almost identical shape-slant invariance for outline targets. Other investigations have produced nearly linear functions also (19, 20, 21).

The curves for the other three conditions (Figures 1b, c, d) appear to depart from linearity, particularly at the upper end, where large slants and small minor/major axis drawing ratios occur. These departures appear to be most pronounced for the coarse-texture condition and least for the untextured surfaces.

The four invariant rates of change pertaining to the direct visual appreciation of the targets can be compared to the conceptualizations provided by Euclidean geometry. For this purpose, the information in Figures 1a, b, c, d, is schematized in Figure 1e. The comparison then consists in noting the nature of these curves in relation to the cosine functions associated with the targets (Figure 1f). *Both sets of curves display orderly relations and both describe the target, one from the observer's point of reference and the other from a geometrical frame of reference.*

The cosine functions all indicate decreases in minor/major axis ratios as inclination is increased. So do the shape-slant curves. Despite a basic similarity between the functions depicted, they are somewhat different. *Five* separate sets of invariant rates of changes (curves) occur from the geometric standpoint, one for each ellipse and one for the circle, surface features of the targets making no difference. In contrast, *four* separate invariant rates of change are shown by the response curves. These four curves do not seem to be determined by the constellation of variables attributed to the target

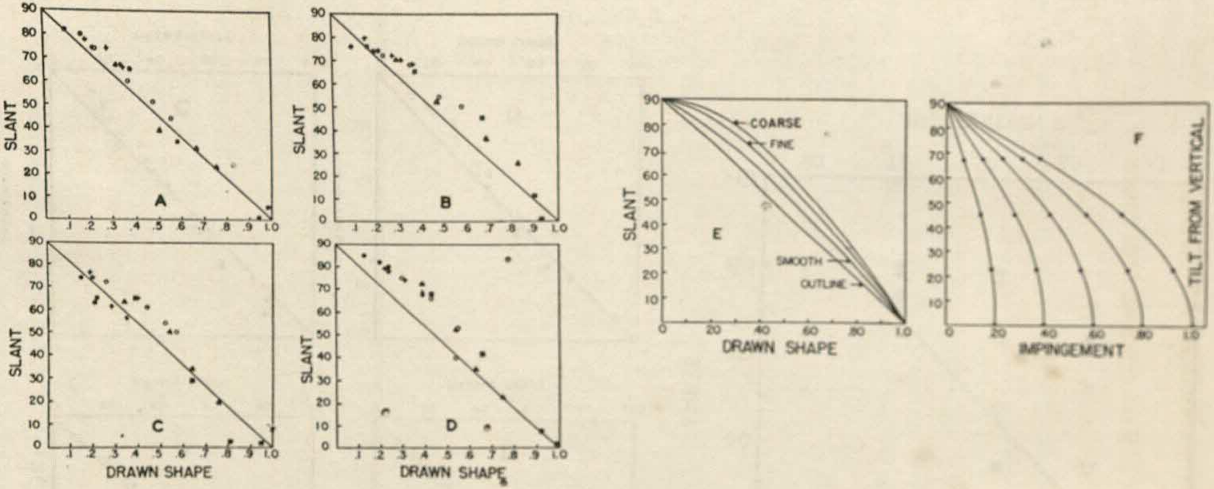


FIGURE 1

The relation between shape index and slant. A is for outline targets; B for untextured surface targets; C for fine texture; and D for coarse texture. Points marked by solid circles = 1 × 5 ellipses; those marked by crosses = 2 × 5 ellipses; open circles = 3 × 5 ellipses; triangle = 4 × 5; and solid square = 5 × 5 inch (circle) target. E and F are schematized relations of targets in the previous graphs. E shows actual results; F, the cosine functions of targets themselves, for various degrees of inclination.

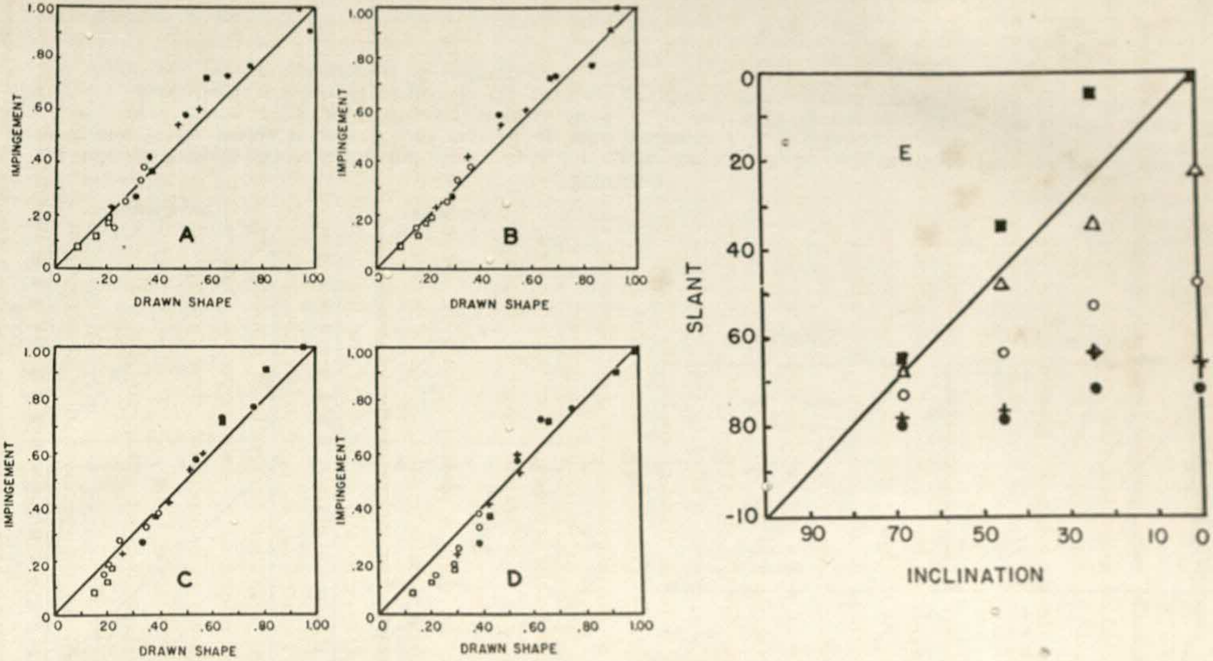


FIGURE 2

Relation of shape responses to frontal plane projections. Curve A = outline targets; B = untextured surface targets; C = fine textured targets; D = coarse textured targets; E = the relation of slant to inclination. Points designated by open squares = 1 × 5 ellipses; open circles = 2 × 5; crosses = 3 × 5; solid circles = 4 × 5; and solid squares = 5 × 5.

by Euclidean geometry, but instead are greatly influenced by their surface characteristics.

Let us examine some possible ways of accounting for invariance. In Figure 2, means of shape and slant responses are related to various physical conditions antecedent to response. In Figures 2a, b, c, d, the minor/major axis ratios of the drawings are plotted against the minor/major axis ratios of the retinal images. In all four cases, it is apparent that characteristics of response are closely tied to the characteristics of the frontal plane projection. The relationship is not quite perfect, however, since the points are not distributed evenly on both sides of the theoretical function over all portions of the graph. At the lower end, the response indexes are ordinarily somewhat greater than those of the image and somewhat less in the middle ranges. This distribution tallies very closely with data Nelson and Bartley (21) reported several years ago and hence is not an artifact of this response sample or experimental set-up. Figures 2a, b, c, d show that the concept of "regression toward the real object" is inadequate to explain the results obtained. If shape were related to the geometrical description of the target, all values lie *between* those given for the retinal image and those given for the "real object." This would demand that all points fall below the curves (diagonals) in the diagrams. Any point appearing *above* the curves indicates what would have to be called "regression *away from* the real object." Therefore, our results are inconsistent with the view expressed in "regression" and "real object" for reasons given by Bartley (2), Nelson (19), and Nelson and Bartley (21).

In Figure 2e, slant is plotted against the inclination of the target, disregarding surface characteristics. Lack of correspondence is clearly evident for every target. For inclination to be considered a direct determinant, all points would have to fall on or near the line indicated. Clearly this does not happen. Haan and Bartley (15) previously showed the same thing.

Whereas Figures 2a, b, c, d relate the minor/major axis ratio of the frontal plane projection to means of drawing responses, Figure 3 relates frontal plane projection to slant responses. Since slant and shape responses have been shown to be invariant, these curves depict what could have been inferred from Figures 1 and 2a, b, c, d, that is, invariance between these changes is definite but not linear, and the outline condition seems to fit the linear possibility best.

To account for the specific nature of shape-slant relationships one might have to resort to contour processes[3] or artifacts in making the response itself,

---

[3] It is necessary to have a self-consistent descriptive terminology for environmental,

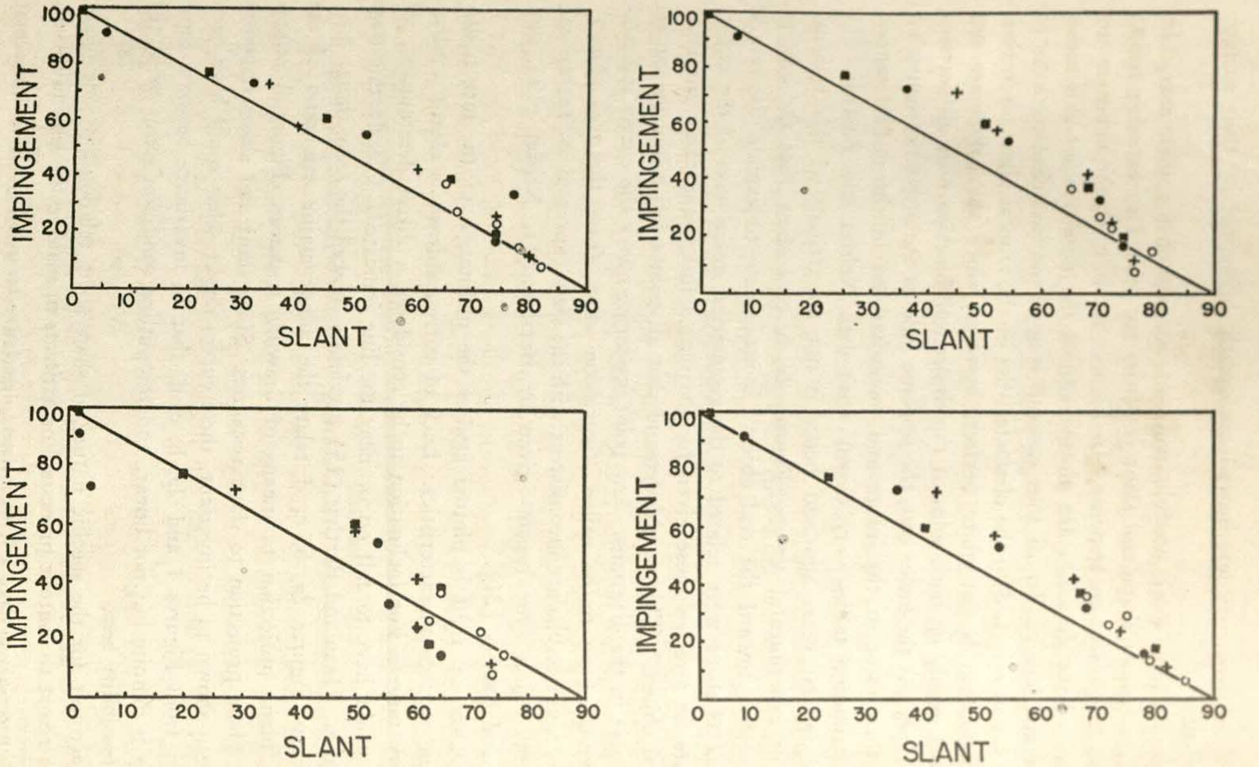


FIGURE 3

Relation of frontal plane projection to slant. Upper left = outline targets; upper right = untextured surface targets; lower left = fine textured targets; lower right = coarse textured targets. Points marked by open circles = 67.5 degrees of target inclination; cross = 45 degrees; solid circles = 22.5 degrees; and solid squares = 0 degrees of inclination.

i.e., drawing per se. Certainly, contour processes may change their character as an accelerating function of distance (1, 9, 10, 11). On the other hand, features such as "perspective," flattening of the near edge, thickening at the points of rotation seem ruled out as major variables. This is said because the small minor/major frontal plane projection values produced by 1 × 5 and 2 × 5 inch targets in the frontal plane act similarly to small minor/major projection values produced by inclined targets of larger physical dimension. This is not what one would expect if these variables were important determiners. This lack of importance seems not to hold for variables of "perspective" (5, 6, 7) placed in more complex environments. Preliminary work of the senior author tends to rule out drawing artifacts as a factor.

The coarse-textured condition produced results unlike those for the outline condition. All slant values are above those predicted by the proportionality of the frontal plane projection and almost all drawing values are greater than those predicted by their respective frontal plane projections. Apparently, a coarsely textured surface does make a difference, although perhaps not the one expected. One might anticipate that texture would affect perceived shape only insofar as it would produce a continuous gradient of texture when inclined. This is suggested by several experimenters (12, 14). Texture gradients also may be presumed to play a role because the portions of the curve in which the greatest deviations from simple linearity occur are associated with steep texture gradients. The effects of this factor can be overrated, however. In Figures 2a, b, c, d, even though inclinations of 45° produce marked gradients in texture, one can see that drawing indexes are sometimes smaller than those for the frontal plane projection, and, conversely, in one case an inclination of 0° gives a larger drawing index than does the frontal plane projection. Hence, one cannot discard the possibility of central determination. Unlike the outline condition, textured surfaces are interposed between the edges and it is possible that texture within borders acts upon contour processes irrespective of gradients per se.

The results for the surface and fine-texture conditions seem to be inter-

---

stimulus, neurophysiological, and perceptual variables if they are to be clearly distinguished and related to one another (25). We shall follow Bartley (2) in his use of *border* to specify the bounds of the target, *edge* to talk about the bounds of what is seen, and *contour* processes to refer to the neurophysiological processes that underlie the seeing of edges. We will use *light gradients* to refer to the response counterpart of a photic gradient, and *surface, texture* or *edge* to refer to perceptual responses related to the specific form of the photic gradient. We shall continue to use the words *target* and *areal* as generic terms and *area* and *photic gradients* as variables specifiable by physical measurement.

mediate between those of outline and coarse texture. The only exception to this occurs over a portion of the fine-texture condition curve.

D. Conclusions

The theoretical approach presented in this paper is in contrast to conventional approaches. We have suggested that previous failures to confirm hypothetical shape-slant invariance have occurred, to a large extent, because of stereotyped uncritical acceptance of some very basic notions concerning the relations between formally developed conceptions of space and human discrimination.

In particular, it is erroneously believed that one can meaningfully compare shape responses and physical indexes such as distal (real) form and proximal (stimulus) form without taking both slant and inclination into account. All constancy indexes presuppose this and it would seem that all are in error because of this.

We have not found it necessary to make the responses stimulus-bound nor to make a special hypothesis for the results, e.g., to relegate the responses to the "visual field" as opposed to the "visual world" as Gibson (12) presumably would do. Neither do our invariant response relations depend upon special mathematical formulation nor any formal interpretation of space for their validity. They are *entirely response-produced functions.* As such they are *purely biological data* and therefore one may justifiably suppose them to be directly dependent upon underlying biological mechanisms (22). No one could claim this for other approaches to invariance, i.e., those making assumptions about the nature of the "object-in-space."

Several factors in addition to the familiar texture gradient of Gibson (12, 13) must be taken into account to explain shape-slant relations. First, our responses were produced in an undifferentiated environment, a condition which Gibson has specifically attempted to divest of any significance in understanding perception in ordinary situations. Second, even extending Gibson's theory of stimulus determination to textured shapes in an undifferentiated field, one would still have to account for the invariant relation occurring when outline targets are used. To say the least, there is a large gap between the nature of response and his psychophysical theory.

In closing, stipulation of antecedent factors as either photic gradients or contour processes in the optic pathway has great advantages over the more typical psychological approaches to perceptual theory. Unlike explanations made by combining veridical statements and cues, photic gradients and contour processes are not of the same level as the phenomena to be accounted for, i.e., responses. Variables of contour process permit discussion of perception in terms of concrete items of neurophysiological activity.

REFERENCES

1. BARTLEY, S. H. Vision: A Study of its Basis. New York: D. Van Nostrand, 1941.
2. ———. Principles of Perception. New York: Harper, 1958.
3. BECK, J., & GIBSON, J. J. The relation of apparent shape to apparent slant in the perception of objects. *J. Exp. Psychol.*, 1955, **50**, 125-133.
4. BRUNSWIK, E. Perception and the representative design of psychological experiments. Berkeley, Calif.: Univ. Calif. Press, 1956.
5. CLARK, W. C., SMITH, A. H., & RABE, A. Retinal gradient of outline as a stimulus for slant. *Canad. J. Psychol.*, 1955, **9**, 247-253.
6. ———. The interaction of surface texture, outline gradient, and ground in the perception of slant. *Canad. J. Psychol.*, 1956, **10**, 1-8.
7. ———. Retinal gradients of outline distortion and binocular disparity as stimuli for slant. *Canad. J. Psychol.*, 1956, **10**, 77-81.
8. CLIFFORD, W. K. The postulates of the science of space. *In* J. R. Newman (Ed.), *The World of Mathematics*. Vol. I. New York: Simon and Schuster, 1956. Pp. 552-567.
9. FRY, G. A. Depression of activity aroused by a flash of light by applying a second flash immediately afterwards to adjacent areas of the retina. *Amer. J. Physiol.*, 1934, **108**, 701-707.
10. ———. Visual perception of space. *Amer. J. Optom.*, 1950, **27**, 531-553.
11. FRY, G. A., & BARTLEY, S. H. The effect of one border in the visual field upon the threshold of another. *Amer. J. Physiol.*, 1935, **112**, 414-421.
12. GIBSON, J. J. The Perception of the Visual World. Boston: Houghton Mifflin & Co., 1950.
13. ———. Perception as a function of stimulation. *In* Koch, S., (Ed.), *Psychology: A Study of a Science. Study I. Conceptual and Systematic*. Vol. I. New York: McGraw-Hill Book Co., Inc., 1959. Pp. 465-502.
14. GRUBER, H. E., & CLARK, W. C. Perception of slanted surfaces. *Percept. and Mot. Skills*, 1956, **6**, 97-106.
15. HAAN, E. L., & BARTLEY, S. H. The apparent orientation of a luminous figure in darkness. *Amer. J. Psychol.*, 1954, **67**, 500-508.
16. HELMHOLTZ, H. The origin and meaning of geometric axioms. *Mind*, 1876, **1**, 3-1-321; 1878, **3**, 212-225.
17. KOFFKA, K. Principles of Gestalt Psychology. New York: Harcourt Brace & Co., Inc., 1935.
18. LANGDON, J. Further studies in the perception of a changing shape. *Quart. J. Exp. Psychol.*, 1953, **5**, 89-107.
19. NELSON, T. M. The perception of a form in a dark field as indicated by the observer's drawings. Unpublished M.A. thesis, Michigan State University, 1953.
20. ———. The relationship of the intrinsic geometric properties of an object to the apparent shape. *Amer. J. Optom.*, 1954, **31**, 289-296.
21. NELSON, T. M., & BARTLEY, S. H. The perception of form in an unstructured field. *J. Gen. Psychol.*, 1956, **54**, 57-63.
22. ———. Numerosity, number, arithmetization, measurement, and psychology. *Phil. Sci.*, 1961, **28**, 178-202.

23. Orbison, W. D. Shape as a function of the vector-field. *Amer. J. Psychol.,* 1939, **52**, 31-45.

24. ———. The correction of an omission in shape as a function of the vector-field. *Amer. J. Psychol.,* 1939, **52**, 309.

25. Rokeach, M., & Bartley, S. H. Some pitfalls in psychological theorizing. *Amer. Psychologist,* 1958, **13**, 283-284.

26. Stavrianos, B. K. The relation of shape perception to explicit judgments of inclination. *Arch. Psychol.,* New York, 1945, No. 296.

27. Thouless, R. H. Phenomenal regression to the real object. *Brit. J. Psychol.,* 1931, Part I, **21**, 339-359; Part II, **22**, 1-30.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology*, 1961, **52**, 491-503.

FANTASY TESTING FOR ESP IN A FOURTH AND FIFTH GRADE CLASS*[1]

Biophysics Department, University of Pittsburgh

M. L. ANDERSON AND R. A. McCONNELL

Experiments of several kinds carried out over the past ten years have indicated that, under certain conditions, extrasensory perception occurs in the classroom. Mr. J. G. van Busschbach, Inspector of Schools, Amsterdam, has conducted statistically significant ESP card-guessing experiments with first, second, fifth, and sixth grade pupils in the Netherlands and America (10, 11, 12, 13, 14). In these experiments the classroom teacher has served as the "sender," and a test leader has acted as liaison between the teacher and the children.

Another series of classroom experiments differed from the van Busschbach tests by suggesting that ESP scoring is associated with the attitude of student and teacher toward each other (1, 4, 5, 6, 7, 8). In this series the highest rate of ESP scoring was attained by students who had a mutually favorable attitude with their teacher. Conversely, the lowest rate of scoring occurred for those students where mutually unfavorable pupil-teacher attitudes prevailed. In this series as well as in the van Busschbach experiments, a procedure was followed wherein the students of a class were tested for ESP in a limited number of guesses within a short interval of time and usually at one session only.

Still another kind of classroom investigation was reported in 1959 (3). Its purpose was to ascertain whether ESP tests conducted with the same students repeatedly and over a long period of time—in this instance, two years—would be successful when there was good reason to expect that positive interpersonal attitudes would predominate. This program of testing was not presented to the students as an experiment; rather, it was presented as a "project" closely tied in with the youngsters' age and stage in school. The form of the experiment was developed within a framework of purposeful fantasy in which it was felt that they would be challenged to make their best effort to score. During the second year, for example, the experiment was designed around a musical theme since the testing was to be carried out during the students' music period. The children were given the name "The Orchestra of Tomorrow" and they were "to code the music of outer space." The youngsters completed 50 symbol guesses per session at a rate of two sessions per month over a six months' period. The response sheets were decoratively designed in keeping with the general theme of music and outer space. The results of this study indicated that a group of elementary school pupils could be so motivated that statistically significant evidence of ESP success would be sustained over a period of several months.

---

* Received in the Editorial Office on August 2, 1961, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This paper is indexed as Publication No. 88 in the Biophysics Department of the University of Pittsburgh, and was prepared under a research program financially supported by the A. W. Mellon Educational and Charitable Trust. The authors wish to acknowledge Mrs. Alberta Covert, teacher in the Verona Elementary School, Verona, Pennsylvania as co-experimenter. Thanks are due to Mr. H. C. Douds, supervising principal of the Verona schools, to the children who cooperated in this experiment, and to Mrs. R. J. Snowdon for assistance in the reading of data.

A. GENESIS

In the spring of 1959, Mrs. Alberta Covert, a teacher in the elementary school of the Pittsburgh suburban community of Verona, wrote MLA of her interest in ESP and the published school experiments. Following an exchange of correspondence and a subsequent move to Pittsburgh, MLA met with Mrs. Covert in early September. It appeared that the Verona school presented an opportunity to design and carry out an experiment in terms of the students' imagination and purposeful fantasy in the manner of MLA's earlier study (3). With Mrs. Covert's enthusiastic approval, the experiment reported here was carried out between November, 1959, and March, 1960, with the 28 pupils of her combined fourth and fifth grade class.

In September, 1959, MLA met with this class. From pictures and models of rockets and missiles about the classroom, it was evident that these young people were interested in the space age. They discussed various rocket and missile projects underway, and MLA then asked if they would like to launch a rocket "mentally." They indicated enthusiastically that they would like to do so.

The group was given the name "Vastronauts" and the rocket to be "launched" was called the Veroco—a combination of "vero" from the name of the school, Verona; and "co" from the teacher's name, Covert. A drawing of the Veroco was designed and a copy of it distributed to each of the students.

B. PROCEDURES

1. *Data Pattern*

The experiment was carried out in three parts, representing Stage I, Stage II, and the Nose Cone of a rocket. Stage I and Stage II of the experiment were tests for clairvoyance (concealed targets), while the Nose Cone was a precognition test (targets still to be selected). MLA's office in the Biophysics Department on the University of Pittsburgh campus was the "control tower." The students were told that in order that there be no interference in their space flight they would receive (for Stages 1 and II only) a sealed individual secret code (the target sheets) prepared at the control tower. In order to launch the Veroco missile and keep it on its course, they would each have to "read" and record their own individual code. After they had done this, Mrs. Covert would send their individually recorded codes (response sheets) to MLA at the control tower, and, after all the codes were computed, MLA would relay to them how the missile was going.

Each of the three parts of the experiment was further subdivided. Stage I consisted of three separate sessions in which 40, 20, and 40 trials or guesses were made respectively by each student. Stage II also consisted of three separate sessions in which 40 trials were made at each session. The Nose Cone was completed in two separate sessions in each of which 20 trials were made.

The response sheet for Stage I, completed at three separate sessions, was designed to correspond with the lower part of the rocket model which the youngsters had received previously. It consisted of five "rocket tubes" (vertical columns) each horizontally divided into 20 blank spaces. Columns 1 and 2 were used at session 1, column 3 at session 2, and columns 4 and 5 at session 3.

The response sheet for Stage II, completed at three separate sessions, corresponded to the cylindrical center portion of the Veroco rocket model. It contained six vertical columns, each horizontally divided into 20 blank spaces. Columns 1 and 2 were used for session one, columns 3 and 4 for session two, and columns 5 and 6 for the third and final session.

The Nose Cone, completed in two sessions, was triangular in shape. It contained 40 blank spaces, pyramided from the bottom to the top. A bold black line separated the bottom 20 trials, which were completed at session one, and the top 20 trials completed at session two.

2. *Target Preparation*

The target symbols for Stages I and II were derived by MLA from the Kendall and Smith's random number table (1951) by listing a suitable number of digits, using a separate table entry for each student and each Stage. Thus, here (as well as in the Nose Cone) a different target order was used for every student and every session. The corresponding target symbols were written directly onto serially numbered target sheets which were identical in form with the response sheets. Two carbon copies of each target sheet were kept in locked files in the Biophysics Department, one by RAM and the other by MLA, while the original was used as described below.

The target symbols for the Nose Cone (precognition test) were made up by MLA after each of the two sessions as follows: Upon request, Mrs. Frances Greene of the Duke University Parapsychology Laboratory—as per previous agreement with the staff there—determined by a prescribed procedure (9) a target entry point for the Kendall and Smith random number table. One record of the entry point was sent to MLA in Pittsburgh and another kept on file at the Duke Laboratory. Starting from the entry point and proceeding in one extended sequence, MLA determined the precognition target symbol order for the session, recording the same with carbon duplicate on sheets serially numbered to correspond to the identification numbers on the students' response sheets. RAM was given the target sheet original and MLA retained the carbon.

3. *Test Materials* (*Response-Sheet Assemblies*)

For Stage I, and likewise for Stage II, each student received a sealed, numbered, manila envelope prepared by MLA, inside of which was a target sheet original and outside of which were the related response sheets. The target sheet was placed between blank paper and opaque aluminum foil, the combination was slipped into the manila envelope, and the flap was closed with two rivet seals. Three blank response sheets were stapled at the top to the outside of the envelope and two carbon papers interleaved so that a copy of each student's ESP calls could be obtained without delay, as described below. The envelope assembly was tested for invisibility of the target symbols by transmitted and reflected light.[2]

---

[2] The Anderson and Anderson-White school experiments (1, 2, 3, 4, 5, 6, 7, 8) for clairvoyance have utilized a similar sealed manila envelope technique. However, the inside target sheet in those experiments was placed between sheets of

For the precognition part of the experiment, since there were no targets available at the time of the test, the test material consisted of two blank response sheets picturing the Nose Cone, stapled together with carbon paper between.

In keeping with the fantasy theme of launching a rocket, the symbols used for session one of Stage I were the cardinal numbers 1 through 5. The youngsters were told that session one was the "count-down" and that each of them had to record the right code contained in his own sealed envelope in order to launch the Veroco.

The symbols for session two consisted of the five arithmetical symbols, + — × = ÷. The youngsters were told that, once having launched the rocket, there would necessarily be a great deal of computational work to keep the electronic equipment going.

For session three the symbols consisted of five colors: red, yellow, blue, green, and orange. The children were asked to imagine that in front of each of them was a control board made up of five colored lights. These lights would flash on and off according to the secret code in their sealed envelopes, and they were to record that code. They were to write the first letter of the color rather than the whole word.

The symbols for session one of Stage II consisted of the letters a, b, x, y, z; for session two, the numbers 1 through 5 were again used; and for session three, the five colors used previously.

For the Nose Cone, or precognition test, the youngsters were given five "space" symbols to use. A moon, a "planet," an "asteroid," a comet, and a star. They were told that since they were now "in orbit," they would have to inform the control tower, by code, what they were seeing. They were to assume that each was at a different window in the rocket. When all the codes reached the control tower, a picture of outer space could then be made from their information.

---

dark poster paper and plain white paper. In the planning of the present experiment, it was found that if an envelope, similarly prepared, were placed before a strong light, the target symbols were discernible. The student subjects in those earlier tests had no opportunity to use this procedure, because the materials were in their hands only during the class hour in which the test was conducted. The fact remains, however, that ordinary sensory knowledge of targets might have been obtained by the experimenters and test leaders in that earlier work without opening the target envelopes. It should be noted, however, that this question cannot be raised concerning the (precognitive) work during the second year of the earlier investigation (3, Tables 6 and 11), whose method is partially replicated in the present research.

4. *Data Gathering*

The tests were administered by the teacher, Mrs. Covert, without the presence of MLA or RAM. The test materials were distributed to the students at the beginning of each session, collected at the end, and stored in a locked drawer when not in use. Students wrote their names on each response-sheet assembly the first time it came to them.

At each session of Stages I and II each pupil wrote in the appropriate spaces of his response sheet the order of the code symbols as he thought they appeared on that particular portion of the identically arranged target sheet within the attached envelope. In the case of the Nose Cone sessions the students were asked to use the pre-chosen set of symbols to record "what each saw from his rocket window." Each student was allowed to make his responses in his own way at his own speed.

At the end of the first and second sessions of Stages I and II and of the first session of the Nose Cone, the teacher removed the bottom response sheets, which were carbon copies of the student responses appearing on the topmost or original response sheets, and mailed these carbons to MLA for checking.

The last session of Stage I was completed the day before a Christmas party to which MLA was invited. MLA accepted the remaining Stage I test materials in a large sealed envelope at that party and delivered the envelope to RAM. The final session test materials for Stage II and the Nose Cone were mailed to the University and received by RAM.

5. *Scoring*

All trials were scored independently by MLA and RAM and discrepancies resolved jointly as described below. All scoring, checking, and correcting on target and response sheets was done using red pencil only. Scoring errors were crossed out rather than erased.

Except for the last session of each part of the experiment, MLA scored the response carbons against her target carbons. For the last session of Stages I and II and the Nose Cone (there being no response carbons), MLA received the original response sheets from RAM, scored these against her target carbons in his presence and returned the unmarked response originals to RAM. These unverified student scores were sent immediately to Mrs. Covert who read them to the class.

The student target envelopes were inspected by RAM for tampering, slit by him at the bottom to remove the target sheet originals, and stored for possible later re-examination.

RAM scored response-sheet originals against target-sheet originals after MLA's scoring as follows. From the response originals, R. J. Snowdon called the responses, determining the symbol in doubtful cases without consulting RAM. RAM checkmarked all hits on the target-sheet original and later totaled the hits for each column (or group) of 20 trials. Blank responses were simulated by RAM from a random-number-table sequence.

RAM and MLA compared scorings as follows: MLA called her hit totals per 20 trials while RAM marked his corresponding totals as right or wrong. Discrepancies were resolved by RAM and MLA jointly.

Tables 1 and 2 and their tests of statistical significance were carried out independently by MLA and RAM from their independent tabulations and discrepancies were resolved by independent repetition. Table 3 was prepared by RAM alone.

C. Results

In the eight sessions of this symbol-guessing experiment twenty-eight student subjects made a total of 6620 binomial trials. As shown in Table 1, the total deviation from mathematical expectation was + 72 with a chance probability for so extreme a score of .028. When evaluated for non-chance behavior by the chi-square of the deviations from expectation in each of the several sessions, the chance probability is .011.

The last two sessions were a test for precognition and yielded a non-significant negative deviation of —8. When the preceding first six sessions, which tested for clairvoyance, are considered separately, the chance probability of their total mean score is .008 and of the distribution of session means is .004.

When the first three clairvoyant sessions constituting Stage I of the rocket fantasy, are considered separately, their total mean score has a P = .006 and their session-mean chi-square has a P = .0004. Clearly, the data are to be regarded as extra-chance.

There were no statistically significant differences between boys and girls or between fourth and fifth grade. The details are shown in Table 2.

An approximate analysis of variance with these non-orthogonal data was made, ignoring sex, grade level, and column scores. The results, shown in Table 3, suggest that there are fluctuations in ESP performance from student-to-student in their total scores and also from day-to-day for the classroom as a whole.

TABLE 1
ESP TOTAL SCORES BY SESSIONS

			Score		Score	Total Devia.		Session Deviations		
	Session	Trials	Total	Devia.	Average†	CR‡	P	χ^2	df	P
"Stage I" (Clairvoyance)										
	1	1080	215	— 1	3.98					
	2	520	139	+35	5.35	3.78	.0002			
	3	920	206	+22	4.48	1.77				
Stage total		2520	560	+56	4.44	2.76	.006	18.0	3	.0004
"Stage II" (Clairvoyance)										
	1	1000	205	+ 5	4.10					
	2	960	205	+13	4.27					
	3	1080	222	+ 6	4.11					
Stage total		3040	632	+24	4.16	1.07				
Clairvoyance total		5560	1192	+80	4.29	2.67	.008	19.5	6	.003
"Nose Cone" (Precognition)										
	1	520	102	— 2	3.92					
	2	540	102	— 6	3.78					
Precognition total		1060	204	— 8	3.85					
Experimental total		6620	1396	+72	4.22	2.20	.028	19.9	8	.011

† Per group of 20 trials. Chance expectation = 4.00.
‡ Yates Correction applied.

TABLE 2
Relationship of ESP Scoring to Grade Level and Sex

	Boys					Girls				
	Boys	Trials	Score Total	Devia.	Score Average†	Girls	Trials	Score Total	Devia.	Score Average†
Fourth grade										
Clairvoyance	4	760	167	+15		9	1780	391	+35	
Precognition	4	160	39	+ 7		9	300	53	— 7	
Total	4	920	206	+22	4.48	9	2080	444	+28	4.27
Fifth grade										
Clairvoyance	4	880	190	+14		11	2140	444	+16	
Precognition	4	160	30	— 2		11	440	82	— 6	
Total	4	1040	220	+12	4.23	11	2580	526	+10	4.08
Experiment total	8	1960	426	+34	4.35	20	4660	970	+38	4.16

† Per group of 20 trials. Chance expectation = 4.00.

Variation among stages is not significantly greater than variation among sessions. Hence, no conclusion can be drawn concerning the nature of the decline from stage-to-stage shown by the score averages 4.44, 4.16, 3.85 (chance expectation = 4.00).

TABLE 3
APPROXIMATE ANALYSIS OF VARIANCE

	Mean Square	df	P
Binomial theory	.160	inf.	—
Residual	.180	170	—
Students	.234	27	~.15
Sessions in stages	.339	5	~.1
Stages	.356	2	—
Students of Session 2, Stage I	.272	25	—

D. DISCUSSION

This experiment represents an extension to a new situation of a previously used (3) ESP-testing technique—that of purposeful fantasy. The data exhibit the anticipated effect to a statistically significant degree and in the absence of a reasonable counterhypothesis might be considered as showing ESP. The deviations from chance expectation are perhaps typical of what is obtained in this kind of research. However, Table 1 shows that they are neither large over-all nor uniformly distributed throughout the experiment. These facts suggest a variety of questions concerning similarities and differences with earlier experiments and concerning existing beliefs about the production of ESP. These questions will be dealt with in a necessarily piecemeal and speculative fashion, taking first the over-all strength of the effect.

As described below, the choice of subjects has utilized the earlier findings of MLA. However, the collaborating teacher, the student subjects, and the laboratory *milieu* were different from those which she previously employed. And in certain respects the experimental controls were more strict than those in her previous work.

A questionnaire showed that the children had a highly favorable attitude toward the teacher, and discussion with the teacher left no doubt as to the warmth of her feeling toward the children. Thus, the group selected for this experiment falls as a whole into the pupil-teacher attitude category that on the basis of earlier research (8) would be expected to show positive ESP scoring.

Similarly, the mean IQ of this class (120) and the mean class grade (4.3

out of a maximum 5) indicate that these children are in the academically deviant group that in earlier work (2) showed ESP scores most deviant from chance expectation.

These observations deserve attention in relating the present experiment to the question of how dependably ESP can be produced in the classroom. In choosing this group the experimenters loaded the experiment as favorably as possible on the basis of previous studies.

What of the variables about which previous studies are silent? Although Mrs. Covert, the test leader, is one of those rare elementary school teachers who establishes rapport with almost all of her pupils, although she showed great and self-generated enthusiasm for the experiment, and although in it she successfully evoked ESP, the authors are inclined to believe that much more than this must ultimately be said concerning the test leader's role in the ESP test situation.

This is not an appropriate place to review the several successful and unsuccessful ESP classroom experiments that point to the importance of this factor. By way of example, however, it may be mentioned that in van Busschbach's fifth and sixth grade studies of USA children (12) where records were kept of the results according to the test leaders of the different classes, consistent unpublished differences among test leaders were found. Indeed, the personality of the person who directly administers the ESP test may be the most important unknown variable in the classroom research.

Although MLA was not present during the test sessions of the present experiment, her role, both in the pupils' thinking and in her own, was a psychologically dominant one. It is a tentative belief, shared by the present authors and by many parapsychologists, that all persons associated with an ESP experiment are, to the degree of their psychological involvement, important to the production (or suppression) of the effect. It is worth pursuing this idea in the light of MLA's five years of successful ESP research.

Superficially, MLA was the one fixed variable when this experiment is compared with her past work. In psychological actuality she is living a different life than when she was at Duke University. The cultural tempo in Pittsburgh is perhaps intermediate between New York City and Boston. By contrast, when one steps off the train in Durham, life slows up by an order-of-magnitude. Beyond that, in Pittsburgh, MLA has returned to graduate study with a heavy scholastic schedule. RAM has observed in her the signs of harassment commonly shown by doctor-of-philosophy candidates. It is encouraging that significantly positive experimental results were obtained despite the new conditions under which she worked.

A second major question concerns the fluctuations of apparent ESP effect within this experiment. The suggestive student-to-student variations in ESP scoring found in the analysis-of-variance are not surprising; the literature abounds with the idea that ESP-demonstrating ability is a highly individual characteristic. The session-to-session scoring variation is of more experimental interest. This kind of fluctuation was also found in MLA's comparable previous work (3, Table 6). From this one may suppose that in the classroom there are important short-term variables which, if understood and controlled, could enhance the ESP effect both in strength and dependability.

Sustained enthusiasm of the immediate participants in an ESP experiment is universally believed by parapsychologists to be of importance when one is working with unselected subjects. Interest and enthusiasm launched the present experiment. Mrs. Covert, the teacher, Mr. Douds, the supervising principal, and the pupil subjects, all were excited with the project. MLA's few visits revealed an animated "stratosphere." The children made a tape recording as a Christmas gift to her which began "Vastronauts, calling Miss Anderson at the control tower." On Valentine's Day she received 28 valentines. Each of the pupils created his own design and verse, and each valentine had something to do with the project. By March the children's parents had become interested in the program, and in early May in a talk to the Parent-Teachers Association, explaining what the project had been, MLA was besieged with questions. The psychological climate throughout the school year was one of interest and excitement.

Thus, the observed session-to-session variation in the ESP scoring cannot be ascribed to a decline or variation in apparent enthusiasm. The experimenters infer that even with a potentially successful combination of teacher, class, and testing method there is psychological happenstance that must be recognized and controlled before ESP can be dependably produced in the classroom.

E. SUMMARY

Using one-in-five binomial target symbols, randomized for each subject, 28 pupils of a fourth-fifth grade class made 6620 ESP guesses in 8 classroom sessions. The total deviation from expectation was +72, for which P = .03. Targets concealed in envelopes attached to the guess-recording forms were used in the first 5560 trials with a deviation of + 80. The remaining 1060 guesses, with deviation —8, were for targets still to be chosen. The largest single-session deviation was + 35 in 520 guesses (P = .0002, uncorrected for selection).

This classroom was chosen for factors favorable to ESP in prior studies. Subjects imagined their guesses to be coded information launching a rocket. The procedure illustrates the kind of "hyper-enthusiastic" approach believed necessary though not sufficient for finding ESP in ordinary subjects. Session variance suggests that even with apparently favorable conditions, ESP cannot be dependably produced in the classroom.

References

1. Anderson, M. L. Clairvoyance and teacher-pupil attitudes in fifth and sixth grades. *J. Parapsychol.*, 1957, **21**, 1-12.
2. ———. The relationship between level of ESP scoring and student class grades. *J. Parapsychol.*, 1959, **23**, 1-18.
3. Anderson, M. L., & Gregory, E. A two-year program of tests for clairvoyance and precognition with a class of public school pupils. *J. Parapsychol.*, 1959, **23**, 149-77.
4. Anderson, M. L., & White, R. Teacher-pupil attitudes and clairvoyance test results. *J. Parapsychol.*, 1956, **20**, 141-57.
5. ———. A further investigation of teacher-pupil attitudes and clairvoyance test results. *J. Parapsychol.*, 1957, **21**, 81-97.
6. ———. ESP score level in relation to students' attitude toward teacher-agents acting simultaneously. *J. Parapsychol.*, 1958, **22**, 20-28.
7. ———. The relationship between changes in student attitude and ESP scoring. *J. Parapsychol.*, 1958, **22**, 167-74.
8. ———. A survey of work on ESP and teacher-pupil attitudes. *J. Parapsychol.*, 1958, **22**, 246-68.
9. Mangan, G. L. Evidence of displacement in a precognition test. *J. Parapsychol.*, 1955, **19**, 35-44.
10. Van Busschbach, J. G. An investigation of extrasensory perception in school children. *J. Parapsychol.*, 1953, **17**, 210-14.
11. ———. A further report on an investigation of ESP in school children. *J. Parapsychol.*, 1955, **19**, 73-81.
12. ———. An investigation of ESP between teachers and pupils in American schools. *J. Parapsychol.*, 1956, **20**, 71-80.
13. ———. An investigation of ESP in the first and second grades of Dutch schools. *J. Parapsychol.*, 1959, **23**, 227-37.
14. ———. An investigation of ESP in first and second grades in American schools. *J. Parapsychol.*, 1961, **25**, (in press).

Biophysics Department
University of Pittsburgh
Pittsburgh 13, Pennsylvania